ON THE QUASI-MONTE CARLO METHOD WITH HALTON POINTS FOR ELLIPTIC PDES WITH LOG-NORMAL DIFFUSION

HELMUT HARBRECHT, MICHAEL PETERS, AND MARKUS SIEBENMORGEN

ABSTRACT. This article is dedicated to the computation of the moments of the solution to elliptic partial differential equations with random, log-normally distributed diffusion coefficients by the quasi-Monte Carlo method. Our main result is that the convergence rate of the quasi-Monte Carlo method based on the Halton sequence for the moment computation depends only linearly on the dimensionality of the stochastic input parameters. Especially, we attain this rather mild dependence on the stochastic dimensionality without any randomization of the quasi-Monte Carlo method under consideration. For the proof of the main result, we require related regularity estimates for the solution and its powers. These estimates are also provided here. Numerical experiments are given to validate the theoretical findings.

1. INTRODUCTION

In this article, we analyze the quasi-Monte Carlo method based on the Halton sequence, cf. [16, 32], to determine the moments of the solution to partial differential equations with random, log-normally distributed diffusion coefficient. Precisely, we consider here equations in divergence form, i.e.

$$-\operatorname{div}\big(a(\mathbf{x},\omega)\nabla u(\mathbf{x},\omega)\big) = f(\mathbf{x}). \tag{1}$$

For simplicity, we impose homogenous boundary conditions.

The efficient treatment of this type of equations has recently been the topic in several works, see e.g. [4, 7, 8, 14, 15, 26, 41]. The method of choice to cope with these equations mainly depends on the quantity of interest. Since the diffusion coefficient is modeled as a stochastic field, it is clear that the solution itself will also be a stochastic field. Therefore, if the solution u itself is of interest, methods like the stochastic Galerkin method, see e.g. [5, 11, 12, 31], or the stochastic collocation method, see e.g. [4, 33], are feasible for its approximation. If one is rather interested in distribution properties of the solution, it might be more convenient to directly approximate the solution's moments, i.e. the expected values of the powers u^p for $p \in \mathbb{N}$. In the latter case, one ends up with a high-dimensional Bochner integration problem which can be solved by quadrature methods. Any quadrature method amounts to the repeated evaluation of the integrand in different sample points or quadrature points. Each such evaluation corresponds to the solution of the equation (1) for a different realization of the parameter ω.

In the present situation of partial differential equations with random, log-normally distributed diffusion coefficient, the domain of integration is unbounded, since the support of the Gaussian density function is the whole real line. Hence, in order to treat the integration problem numerically with a quasi-Monte Carlo method, the domain of integration has to be mapped back to the unit cube. This may introduce singularities in the integrand which complicates the analysis.

To overcome this obstruction, one can randomize the quasi-Monte Carlo method, which leads for example to randomly shifted lattice rules, see e.g. [43]. Randomized quasi-Monte Carlo methods are well suited for the treatment of integration problems on unbounded domains, cf. [15, 28]. In particular, it is shown in [15] that randomly shifted lattice rules converge, independent of the dimensionality, linearly in the number of sample points provided that the input data are smooth enough. This means that they converge twice as fast as the standard Monte Carlo method,

---

2000 *Mathematics Subject Classification.* 65C05, 65C30, 60H25, 65N30.

This research has been supported by the Swiss National Science Foundation (SNSF) through the project "Rapid Solution of Boundary Value Problems on Stochastic Domains".

see e.g. [17]. Compared to deterministic methods like the quasi-Monte Carlo method based on deterministic point sequences or the sparse grid quadrature, randomized lattice rules have the advantage that they require weaker assumptions on the integrand's regularity. As we will see later on, this turns out to hold true also for the proposed quadrature method.

Since randomized quadrature methods provide only stochastic convergence in the mean square sense with respect to the randomization, we aim in the article on hand on a deterministic quadrature method. Then, the error is usually measured in a stronger, but deterministic norm. Moreover, the results of deterministic quadrature methods are always reproducible. Specifically, we analyze the application of the quasi-Monte Carlo quadrature based on Halton points which uses a classical *low discrepancy sequence* and is easy to construct.

The common line of action for the solution of (1) is based on the separation of the stochastic variable and the spatial variable in the diffusion coefficient a. This is achieved by the computation of the so-called Karhunen-Loève expansion, cf. [29], which coincides with a series expansion of a by L^2-orthogonal functions. Thus, the diffusion coefficient depends in principle on infinitely many terms. Depending on the desired accuracy, this series has to be truncated appropriately. This implies that the dimensionality of the integration problem, which is directly coupled to the length of the truncated Karhunen-Loève expansion, increases for higher accuracies. Therefore, it is crucial to construct methods which are as far as possible independent of the length of the Karhunen-Loève expansion. Especially, we want to avoid the exponential dependence of the computational cost on the dimensionality, which is known as the curse of dimensionality. As we will show, the convergence of quasi-Monte Carlo quadrature rules based on the Halton sequence depend only polynomially on the dimensionality. More precisely, the convergence rate, in terms of the number of integration points, depends only linearly on the problem's dimensionality if the Karhunen-Loève expansion decays sufficiently fast. This is the main result of this article. The proof is based on the fact that the Halton sequence avoids the boundary of the integration domain, which has originally been shown in [37].

We like to remark that multilevel techniques, like the multilevel Monte Carlo method, cf. [6, 8, 13, 21, 22], have recently become quite popular to keep the computational cost for the solution of (1) low. However, in [18, 19], it is shown that arbitrary quadrature rules can be accelerated by multilevel techniques, yielding the related multilevel quadrature methods. Especially, faster converging quadrature rules result in a faster converging multilevel quadrature method.

The rest of this article is organized as follows. Section 2 specifies the diffusion problem under consideration and the corresponding framework. In particular, the parametric reformulation as a high-dimensional deterministic problem is performed here. In Section 3, we derive the crucial regularity estimates for the stochastic diffusion problem's solution and its powers. In Section 4, we elaborate the quasi-Monte Carlo quadrature based on the Halton sequence and prove its almost dimension independent convergence. Finally, Section 5 validates the theoretical findings by some basic one-dimensional numerical examples. For more sophisticated examples, we refer to the recent work [18].

In the following, in order to avoid the repeated use of generic but unspecified constants, by $C \lesssim D$ we mean that C can be bounded by a multiple of D, independently of parameters which C and D may depend on. Obviously, $C \gtrsim D$ is defined as $D \lesssim C$, and $C \eqsim D$ as $C \lesssim D$ and $C \gtrsim D$.

2. Problem setting

In the following, let $D \subset \mathbb{R}^d$ for $d \in \mathbb{N}$ be a domain with Lipschitz continuous boundary and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete, separable probability space with σ-field $\mathcal{F} \subset 2^\Omega$ and probability measure $\mathbb{P}$. Let the random function $u(\omega) \in H_0^1(D)$ be the solution to the stochastic diffusion problem

$$-\operatorname{div}\big(a(\omega)\nabla u(\omega)\big) = f \text{ in } D \quad \text{for almost every } \omega \in \Omega \tag{2}$$

with (deterministic) data $f \in L^2(D)$. Instead of directly approximating the probably infinite dimensional solution u itself, we rather intend to compute the solution's moments:

$$\mathcal{M}^p u := \mathbb{E}[u(\cdot,\omega)^p].$$

Especially, the solution's expectation is given by

$$\mathbb{E}_u(\mathbf{x}) = \int_\Omega u(\mathbf{x},\omega)\,\mathrm{d}\mathbb{P}(\omega) \in H_0^1(D) \tag{3}$$

and its variance by

$$\mathbb{V}_u(\mathbf{x}) = \mathbb{E}_{u^2}(\mathbf{x}) - \mathbb{E}_u^2(\mathbf{x}) = \int_\Omega u^2(\mathbf{x},\omega)\,\mathrm{d}\mathbb{P}(\omega) - \mathbb{E}_u^2(\mathbf{x}) \in W_0^{1,1}(D). \tag{4}$$

They correspond to the first and the second (centered) moment of the solution u. As we will show later on, it holds more generally for a sufficiently smooth diffusion coefficient a and data $f \in L^p(D)$ that the moments satisfy $\mathcal{M}^p u \in W_0^{1,1}(D)$. Note that the knowledge of all moments of u is sufficient to determine the related distribution.

We consider here the log-normal situation, where the logarithm of the diffusion coefficient is a centered Gaussian field which can be represented by a Karhunen-Loève expansion, cf. [29],

$$\log\big(a(\mathbf{x},\omega)\big) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\varphi_k(\mathbf{x})\psi_k(\omega). \tag{5}$$

The functions $\{\varphi_k\}_k$ are pairwise orthonormal functions in $L^2(D)$ which are assumed to be bounded and $\{\psi_k\}_k$ are independent, standard normally distributed random variables, i.e. $\psi_k(\omega) \sim \mathcal{N}(0,1)$. For the convergence of the series in (5), we assume that the sequence

$$\gamma_k := \big\|\sqrt{\lambda_k}\varphi_k\big\|_{L^\infty(D)} \tag{6}$$

satisfies $\{\gamma_k\}_k \in \ell^1(\mathbb{N})$, where $\ell^1(\mathbb{N})$ is the space of absolutely summable sequences.

For numerical issues, it is reasonable to assume that the Karhunen-Loève expansion is either finite of length m or it is appropriately truncated after m terms. We will explicitly make use of this assumption in the following. Even so, we explicitly allow that $m \to \infty$ as the accuracy requirements increase. The possibly occurring truncation error for the Karhunen-Loève expansion has been discussed in [7].

The orthogonality of the sequence $\{\psi_k\}_k$ already implies the stochastic independence of this sequence in the Gaussian case. Therefore, the pushforward measure $\mathbb{P}_{\boldsymbol{\psi}} := \mathbb{P} \circ \boldsymbol{\psi}$ with respect to the measurable mapping

$$\boldsymbol{\psi}\colon \Omega \to \mathbb{R}^m, \quad \omega \mapsto \boldsymbol{\psi}(\omega) := \big(\psi_1(\omega), \ldots, \psi_m(\omega)\big)$$

is given by a joint density function with respect to the Lebesgue measure, i.e.

$$\boldsymbol{\rho}(\mathbf{y}) := \prod_{k=1}^{m} \rho(y_k), \quad \text{where} \quad \rho(y) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right). \tag{7}$$

Having this representation at hand, we reformulate the stochastic problem (2) as a parametric deterministic problem. To that end, we substitute the random variables ψ_k by their coordinates $y_k \in \mathbb{R}$. Then, we define the parameterized and truncated diffusion coefficient via

$$a(\mathbf{x},\mathbf{y}) := \exp\left(\sum_{k=1}^{m} \sqrt{\lambda_k}\varphi_k(\mathbf{x}) y_k\right) \tag{8}$$

for all $\mathbf{x} \in D$ and $\mathbf{y} = (y_1, y_2, \ldots, y_m) \in \mathbb{R}^m$. We arrive at the variational formulation for the parametric diffusion problem:

$$\begin{aligned} &\text{find } u \in L^p_{\boldsymbol{\rho}}\big(\mathbb{R}^m; W_0^{1,p}(D)\big) \text{ such that} \\ &\qquad -\operatorname{div}\big(a(\mathbf{x},\mathbf{y})\nabla u(\mathbf{x},\mathbf{y})\big) = f(\mathbf{x}) \text{ in } D \text{ for all } \mathbf{y} \in \mathbb{R}^m. \end{aligned} \tag{9}$$

The parametrization of the problem yields a change of the domain of integration for the computation of the moments. We now integrate with respect to the pushforward measure which induces the integral transform

$$\mathcal{M}^p u = \int_\Omega u^p(\omega)\,\mathrm{d}\mathbb{P}(\omega) = \int_{\mathbb{R}^m} u^p(\mathbf{y})\boldsymbol{\rho}(\mathbf{y})\,\mathrm{d}\mathbf{y}. \tag{10}$$

Here and in the sequel, for a given Banach space X, the Bochner space $L^p_{\boldsymbol{\rho}}(\mathbb{R}^m; X)$, $1 \leqslant p \leqslant \infty$, consists of all equivalence classes of strongly measurable functions $v : \mathbb{R}^m \to X$ whose norm

$$\|v\|_{L^p_{\boldsymbol{\rho}}(\mathbb{R}^m;X)} := \begin{cases} \left(\displaystyle\int_{\mathbb{R}^m} \|v(\cdot, \mathbf{y})\|_X^p \boldsymbol{\rho}(\mathbf{y}) \, \mathrm{d}\mathbf{y} \right)^{1/p}, & p < \infty \\ \operatorname*{ess\,sup}_{\mathbf{y} \in \mathbb{R}^m} \|v(\cdot, \mathbf{y})\|_X, & p = \infty \end{cases}$$

is finite. If $p = 2$ and X is a separable Hilbert space, then the Bochner space is isomorphic to the tensor product space $L^2_{\boldsymbol{\rho}}(\mathbb{R}^m) \otimes X$. Note that, for notational convenience, we will always write $v(\mathbf{x}, \mathbf{y})$ instead of $\big(v(\mathbf{y})\big)(\mathbf{x})$ if $v \in L^p_{\boldsymbol{\rho}}(\mathbb{R}^m; X)$.

The stochastic diffusion coefficient $a(\mathbf{x}, \mathbf{y})$ is neither uniformly bounded away from zero nor uniformly bounded from above with respect to $\mathbf{y}$. Consequently, it is impossible to show unique solvability in the classical way for elliptic boundary value problems. Especially, the Lax-Milgram theorem does not directly apply to the problem (2). Nevertheless, we have for each fixed $\mathbf{y} \in \mathbb{R}^m$ the bounds

$$0 < a_{\min}(\mathbf{y}) := \operatorname*{ess\,inf}_{\mathbf{x} \in D} a(\mathbf{x}, \mathbf{y}) \leqslant \operatorname*{ess\,sup}_{\mathbf{x} \in D} a(\mathbf{x}, \mathbf{y}) =: a_{\max}(\mathbf{y}) < \infty. \tag{11}$$

Obviously, it holds

$$a_{\min}(\mathbf{y}) \geqslant \exp\left(-\sum_{k=1}^m |\gamma_k y_k| \right) \quad \text{and} \quad a_{\max}(\mathbf{y}) \leqslant \exp\left(\sum_{k=1}^m |\gamma_k y_k| \right).$$

Due to (11), for every fixed $\mathbf{y} \in \mathbb{R}^m$, the problem to find $u \in H^1_0(D)$ such that

$$-\operatorname{div}\big(a(\mathbf{x}, \mathbf{y}) \nabla u(\mathbf{x}, \mathbf{y})\big) = f(\mathbf{x}) \text{ in } D \tag{12}$$

is elliptic and admits a unique solution $u(\cdot, \mathbf{y}) \in H^1_0(D)$ which satisfies the stability estimate

$$\|u(\cdot, \mathbf{y})\|_{H^1(D)} \lesssim \frac{1}{a_{\min}(\mathbf{y})} \|f\|_{L^2(D)}. \tag{13}$$

We refer the reader to e.g. [41] for a proof of this result.

Remark 2.1. *Here and in the sequel, the Sobolev space $H^1_0(D)$ is considered to be equipped with the norm*

$$\|\cdot\|_{H^1(D)} := \|\nabla \cdot\|_{[L^2(D)]^d}.$$

Likewise, we use the corresponding norms for the Sobolev spaces $W^{1,p}_0(D)$ for $p \neq 2$, i.e.

$$\|\cdot\|_{W^{1,p}(D)} := \|\nabla \cdot\|_{[L^p(D)]^d}.$$

Since we only consider homogenous Dirichlet problems, by Sobolev's norm equivalence theorem, cf. [2]*, they all induce equivalent norms in comparison to the standard $\|\cdot\|_{W^{1,p}(D)}$-norm*

$$\|v\|_{W^{1,p}(D)} = \left(\|v\|^p_{L^p(D)} + \sum_{k=1}^d \left\| \frac{\partial}{\partial x_k} v \right\|^p_{L^p(D)} \right)^{1/p}$$

for these spaces. Of course, all results are straightforwardly extendable to the case of non-homogenous Dirichlet problems.

3. Regularity of the solution

The topic we address in this article is the computation of the mean and the higher order moments of the solution of (9) by a fully deterministic quadrature rule. Therefore, in order to establish error bounds for the application of quasi-Monte Carlo quadrature rules, we consider in this section the regularity of the solution u and its powers, i.e. u^p for $p \in \mathbb{N}$. This issue has already been discussed for the case of u in [4, 7, 26, 41]. We will compile and augment here some of the results which originate from those articles for our framework.

At first, we shall fix some notation. For a multi-index $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathbb{N}^m$, the related multidimensional derivative is denoted by

$$\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} v(\mathbf{y}) := \frac{\partial^{\alpha_1}}{\partial y_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial y_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_d}}{\partial y_d^{\alpha_d}} v(\mathbf{y}).$$

Furthermore, we set $|\boldsymbol{\alpha}| := \sum_{k=1}^m \alpha_k$ and, for a vector $\boldsymbol{\beta} \in \mathbb{R}^m$, we define $\boldsymbol{\beta}^{\boldsymbol{\alpha}} := \prod_{k=1}^m \beta_k^{\alpha_k}$. Finally, the binomial coefficient for multi-indices is given by

$$\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} = \frac{\boldsymbol{\alpha}!}{\boldsymbol{\beta}!(\boldsymbol{\alpha}-\boldsymbol{\beta})!} = \frac{\alpha_1!\alpha_2!\cdots\alpha_m!}{\beta_1!\cdots\beta_m!(\alpha_1-\beta_1)!\cdots(\alpha_m-\beta_m)!}.$$

The differentiability of u follows straightforwardly from the differentiability of the diffusion coefficient a, cf. [26]. In particular, we shall use the following lemma from [26] which is adjusted for our purposes. Therefore, we denote by $\boldsymbol{\gamma} := (\gamma_1, \gamma_2, \ldots, \gamma_m)$ the first m elements of the sequence in (6).

Lemma 3.1. *For the solution u to (9) and every $\mathbf{y} \in \mathbb{R}^m$, the estimate*

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u(\cdot,\mathbf{y})\|_{H^1(D)} \leqslant |\boldsymbol{\alpha}|! \left(\frac{\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}} \sqrt{\frac{a_{\max}(\mathbf{y})}{a_{\min}(\mathbf{y})}} \|u(\cdot,\mathbf{y})\|_{H^1(D)} \tag{14}$$

holds for all multi-indices $\boldsymbol{\alpha} \in \mathbb{N}^m$.

This result shows the regularity of the solution u. For the regularity of u^2, we have then the following proposition.

Proposition 3.2. *The derivatives of u^2, where u is the solution of (9), satisfy for all multi-indices $\boldsymbol{\alpha} \in \mathbb{N}^m$ and every $\mathbf{y} \in \mathbb{R}^m$ that*

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u^2(\cdot,\mathbf{y})\|_{W^{1,1}(D)} \lesssim (|\boldsymbol{\alpha}|+1)! \left(\frac{\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}} \frac{a_{\max}(\mathbf{y})}{a_{\min}(\mathbf{y})} \|u(\cdot,\mathbf{y})\|_{H^1(D)}^2. \tag{15}$$

Proof. By the Leibniz rule, we obtain

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u^2(\cdot,\mathbf{y})\|_{W^{1,1}(D)} \leqslant \sum_{\boldsymbol{\beta}\leqslant\boldsymbol{\alpha}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} \|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} u(\cdot,\mathbf{y}) \partial_{\mathbf{y}}^{\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{W^{1,1}(D)}. \tag{16}$$

Each of the summands can be estimated as follows

$$\begin{aligned}
&\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} u(\cdot,\mathbf{y}) \partial_{\mathbf{y}}^{\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{W^{1,1}(D)} \\
&\quad = \|\nabla\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} u(\cdot,\mathbf{y}) \partial_{\mathbf{y}}^{\boldsymbol{\beta}} u(\cdot,\mathbf{y}) + \partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} u(\cdot,\mathbf{y}) \nabla\partial_{\mathbf{y}}^{\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{[L^1(D)]^d} \\
&\quad \leqslant \|\nabla\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{[L^2(D)]^d} \|\partial_{\mathbf{y}}^{\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{L^2(D)} + \|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{L^2(D)} \|\nabla\partial_{\mathbf{y}}^{\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{[L^2(D)]^d} \\
&\quad \lesssim \|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{H^1(D)} \|\partial_{\mathbf{y}}^{\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{H^1(D)} + \|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{H^1(D)} \|\partial_{\mathbf{y}}^{\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{H^1(D)}.
\end{aligned}$$

Now, the application of Lemma 3.1 leads to

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} u(\cdot,\mathbf{y}) \partial_{\mathbf{y}}^{\boldsymbol{\beta}} u(\cdot,\mathbf{y})\|_{W^{1,1}(D)} \lesssim |\boldsymbol{\beta}|!\,|(\boldsymbol{\alpha}-\boldsymbol{\beta})|! \left(\frac{\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}} \frac{a_{\max}(\mathbf{y})}{a_{\min}(\mathbf{y})} \|u(\cdot,\mathbf{y})\|_{H^1(D)}^2.$$

By inserting this inequality into (16), we conclude

$$\begin{aligned}
\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u^2(\cdot,\mathbf{y})\|_{W^{1,1}(D)} &\lesssim \sum_{\boldsymbol{\beta}\leqslant\boldsymbol{\alpha}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} |\boldsymbol{\beta}|!\,|(\boldsymbol{\alpha}-\boldsymbol{\beta})|! \left(\frac{\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}} \frac{a_{\max}(\mathbf{y})}{a_{\min}(\mathbf{y})} \|u(\cdot,\mathbf{y})\|_{H^1(D)}^2 \\
&= \left(\frac{\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}} \frac{a_{\max}(\mathbf{y})}{a_{\min}(\mathbf{y})} \|u(\cdot,\mathbf{y})\|_{H^1(D)}^2 \sum_{k=0}^{|\boldsymbol{\alpha}|} (|\boldsymbol{\alpha}|-k)!\,k! \sum_{|\boldsymbol{\beta}|=k} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}.
\end{aligned}$$

In view of

$$\sum_{k=0}^{|\boldsymbol{\alpha}|} (|\boldsymbol{\alpha}|-k)!\,k! \sum_{|\boldsymbol{\beta}|=k} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} = \sum_{k=0}^{|\boldsymbol{\alpha}|} (|\boldsymbol{\alpha}|-k)!\,k! \binom{|\boldsymbol{\alpha}|}{k} = \sum_{k=0}^{|\boldsymbol{\alpha}|} |\boldsymbol{\alpha}|! = (|\boldsymbol{\alpha}|+1)!,$$

we finally arrive at the assertion (15). □

For higher order moments, we need some stronger regularity assumptions on the data f.

Proposition 3.3. *Let D be a domain with sufficiently smooth boundary and let $p > 2$. If the data f satisfies $f \in L^p(D)$, then the solution u to (9) is contained in $W_0^{1,p}(D)$ and meets the stability estimate*

$$\|u(\cdot, \mathbf{y})\|_{W^{1,p}(D)} \lesssim \frac{1}{a_{\min}(\mathbf{y})} \|f\|_{L^p(D)}. \tag{17}$$

Moreover, the derivatives of u with respect to the parametric variable $\mathbf{y}$ can be estimated by

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u(\cdot, \mathbf{y})\|_{W^{1,p}(D)} \leqslant |\boldsymbol{\alpha}|! \left(\frac{C(p, D)\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}} \frac{a_{\max}(\mathbf{y})}{a_{\min}(\mathbf{y})} \|u(\cdot, \mathbf{y})\|_{W^{1,p}(D)} \tag{18}$$

with a constant $C(p, D) > 0$ only dependent on p and the domain D. Additionally, the derivatives of the powers u^p with respect to the parametric variable $\mathbf{y}$ fulfill

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u^p(\cdot, \mathbf{y})\|_{W^{1,1}(D)} \lesssim |\boldsymbol{\alpha}|! \left(\frac{C(p, D)p\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}} \left(\frac{a_{\max}(\mathbf{y})}{a_{\min}(\mathbf{y})}\right)^p \|u(\cdot, \mathbf{y})\|_{W^{1,p}(D)}^p. \tag{19}$$

Proof. At first, we notice that the bilinear form

$$(u, v)_{H_0^1(D)} := (\nabla v, \nabla u)_{L^2(D)}$$

defines an inner product on the Hilbert space $H_0^1(D)$. Let $1 < p, q < \infty$ be dual exponents, i.e. $1/p + 1/q = 1$. It is proven in [42] that for each function $u \in W_0^{1,p}(D)$ the estimate

$$\|u\|_{W^{1,p}(D)} = \|\nabla u\|_{[L^p(D)]^d} = \sup_{0 \neq v \in [L^q(D)]^d} \frac{(\nabla u, v)_{L^2(D)}}{\|v\|_{[L^q(D)]^d}} \leqslant C(p, D) \sup_{0 \neq v \in W_0^{1,q}(D)} \frac{(u, v)_{H_0^1(D)}}{\|v\|_{W^{1,q}(D)}}$$

is valid with a constant $C(p, D) > 1$. This follows from the fact that $W_0^{1,q}(D)$ is densely embedded into $[L^q(D)]^d$ by the mapping $v \mapsto \nabla v$, cf. [42]. From this, we derive

$$\|u(\cdot, \mathbf{y})\|_{W^{1,p}(D)} \leqslant C(p, D) \sup_{0 \neq v \in W_0^{1,q}(D)} \frac{\big(u(\cdot, \mathbf{y}), v\big)_{H_0^1(D)}}{\|v\|_{W^{1,q}(D)}} \leqslant \frac{C(p, D)}{a_{\min}(\mathbf{y})} \sup_{0 \neq v \in W_0^{1,q}(D)} \frac{\mathcal{B}_{\mathbf{y}}(u, v)}{\|v\|_{W^{1,q}(D)}}.$$

Herein, we set

$$\mathcal{B}_{\mathbf{y}}(u, v) := \int_D a(\mathbf{x}, \mathbf{y}) \nabla u(\mathbf{x}, \mathbf{y}) \nabla v(\mathbf{x}) \,\mathrm{d}\mathbf{x}, \tag{20}$$

which is the bilinear form related to the variational formulation of (12) for a fixed value of the parameter $\mathbf{y}$. In full, this variational formulation reads

$$\mathcal{B}_{\mathbf{y}}(u, v) = (f, v)_{L^2(D)} \quad \text{for all } v \in H_0^1(D). \tag{21}$$

From $q < 2$, we infer that $H_0^1(D) \subset W_0^{1,q}(D)$. Since $f \in L^p(D)$, it is easy to verify by a density argument that equation (21) remains valid for $v \in W_0^{1,q}(D)$. Therefore, we have

$$\sup_{0 \neq v \in W_0^{1,q}(D)} \frac{\mathcal{B}_{\mathbf{y}}(u, v)}{\|v\|_{W^{1,q}(D)}} = \sup_{0 \neq v \in W_0^{1,q}(D)} \frac{(f, v)_{L^2(D)}}{\|v\|_{W^{1,q}(D)}} \lesssim \|f\|_{L^p(D)},$$

which follows from the Hölder inequality and the estimate $\|v\|_{L^q(D)} \lesssim \|v\|_{W^{1,q}}$. This establishes the inequality (17).

In the second estimate (18), the estimate (14) which has been proven in [26] is modified in order to bound the derivatives $\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u$ in the $W^{1,p}(D)$-norm. It is shown with arguments similar to those in [26]. We sketch here the essential ideas of the proof which is based on induction. Concretely, we show that

$$\|a(\cdot, \mathbf{y}) \nabla \partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u(\cdot, \mathbf{y})\|_{[L^p(D)]^d} \leqslant |\boldsymbol{\alpha}|! \left(\frac{C(p, D)\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}} \|a(\cdot, \mathbf{y}) \nabla u(\cdot, \mathbf{y})\|_{[L^p(D)]^d}. \tag{22}$$

The case $|\boldsymbol{\alpha}| = 0$ is trivial. For $|\boldsymbol{\alpha}| = k > 0$, we have

$$\|a(\cdot, \mathbf{y}) \nabla \partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u(\cdot, \mathbf{y})\|_{[L^p(D)]^d} \leqslant C(p, D) \sup_{0 \neq v \in W_0^{1,q}(D)} \frac{\mathcal{B}_{\mathbf{y}}\big(\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u(\cdot, \mathbf{y}), v\big)}{\|v\|_{W^{1,q}(D)}}. \tag{23}$$

Now, the differentiation of the bilinear form (20) with respect to $\mathbf{y}$ yields

$$\partial_{\mathbf{y}}^{\boldsymbol{\alpha}}\mathcal{B}_{\mathbf{y}}\big(u(\cdot,\mathbf{y}),v\big) = \mathcal{B}_{\mathbf{y}}\big(\partial_{\mathbf{y}}^{\boldsymbol{\alpha}}u(\cdot,\mathbf{y}),v\big) + \sum_{\mathbf{0}\neq\boldsymbol{\beta}\leqslant\boldsymbol{\alpha}}\int_D \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\partial_{\mathbf{y}}^{\boldsymbol{\beta}}a(\mathbf{x},\mathbf{y})\nabla\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}}u(\mathbf{x},\mathbf{y})\nabla v(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Therefore, from the differentiation of the variational formulation (21), we obtain

$$\begin{aligned}\mathcal{B}_{\mathbf{y}}\big(\partial_{\mathbf{y}}^{\boldsymbol{\alpha}}u(\cdot,\mathbf{y}),v\big) &\leqslant \sum_{\mathbf{0}\neq\boldsymbol{\beta}\leqslant\boldsymbol{\alpha}}\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\bigg\|\frac{\partial_{\mathbf{y}}^{\boldsymbol{\beta}}a(\cdot,\mathbf{y})}{a(\cdot,\mathbf{y})}\bigg\|_{L^\infty(D)}\int_D \big|a(\mathbf{x},\mathbf{y})\nabla\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}}u(\mathbf{x},\mathbf{y})\nabla v(\mathbf{x})\big|\,\mathrm{d}\mathbf{x}\\ &\leqslant \sum_{\mathbf{0}\neq\boldsymbol{\beta}\leqslant\boldsymbol{\alpha}}\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\boldsymbol{\gamma}^{\boldsymbol{\beta}}\big\|a(\cdot,\mathbf{y})\nabla\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}}u(\cdot,\mathbf{y})\big\|_{[L^p(D)]^d}\|v\|_{W^{1,q}(D)}.\end{aligned}$$

Inserting this into (23) leads to

$$\big\|a(\cdot,\mathbf{y})\nabla\partial_{\mathbf{y}}^{\boldsymbol{\alpha}}u(\cdot,\mathbf{y})\big\|_{[L^p(D)]^d} \leqslant C(p,D)\sum_{\mathbf{0}\neq\boldsymbol{\beta}\leqslant\boldsymbol{\alpha}}\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\boldsymbol{\gamma}^{\boldsymbol{\beta}}\big\|a(\cdot,\mathbf{y})\nabla\partial_{\mathbf{y}}^{\boldsymbol{\alpha}-\boldsymbol{\beta}}u(\cdot,\mathbf{y})\big\|_{[L^p(D)]^d}.$$

The inequality (22) follows then by inserting the induction hypothesis and some combinatorial estimates as in [26].

Finally, to establish estimate (19), we apply Faà di Bruno's formula, cf. [9]. For $n := |\boldsymbol{\alpha}|$, this formula provides

$$\partial_{\mathbf{y}}^{\boldsymbol{\alpha}}u^p(\cdot,\mathbf{y}) = \sum_{r=1}^{n} p(p-1)\cdots(p-r+1)u^{p-r}(\cdot,\mathbf{y})\sum_{P(\boldsymbol{\alpha},r)}\boldsymbol{\alpha}!\prod_{j=1}^{n}\frac{\big(\partial_{\mathbf{y}}^{\boldsymbol{\beta}_j}u(\cdot,\mathbf{y})\big)^{k_j}}{k_j!\boldsymbol{\beta}_j!}, \tag{24}$$

Here, the set $P(\boldsymbol{\alpha},r)$ contains restricted integer partitions of a multi-index $\boldsymbol{\alpha}$ into r non-vanishing multi-indices. It is defined according to

$$\begin{aligned}P(\boldsymbol{\alpha},r) := \bigg\{\big((k_1,\boldsymbol{\beta}_1),\ldots,(k_n,\boldsymbol{\beta}_n)\big)\in(\mathbb{N}\times\mathbb{N}^m)^n : \sum_{i=1}^{n}k_i\boldsymbol{\beta}_i = \boldsymbol{\alpha}, \sum_{i=1}^{n}k_i = r,\\ \text{and } \exists\, 1\leqslant s\leqslant n \text{ such that } k_i = 0 \text{ and } \boldsymbol{\beta}_i = \mathbf{0} \text{ for all } 1\leqslant i\leqslant n-s,\\ k_i > 0 \text{ for all } n-s+1\leqslant i\leqslant n \text{ and } \mathbf{0}\prec\boldsymbol{\beta}_{n-s+1}\prec\cdots\prec\boldsymbol{\beta}_n\bigg\}.\end{aligned}$$

Herein, for multi-indices $\boldsymbol{\beta},\boldsymbol{\beta}'\in\mathbb{N}^m$, the relation $\boldsymbol{\beta}\prec\boldsymbol{\beta}'$ means either $|\boldsymbol{\beta}|<|\boldsymbol{\beta}'|$ or, if $|\boldsymbol{\beta}|=|\boldsymbol{\beta}'|$, it denotes the lexicographical order which means that $\beta_1=\beta_1',\ldots,\beta_k=\beta_k'$ and $\beta_{k+1}<\beta_{k+1}'$ for some $0\leqslant k<m$.

Equation (24) together with the generalized Hölder inequality yields

$$\begin{aligned}&\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}}u^p(\cdot,\mathbf{y})\|_{W^{1,1}(D)}\\ &\leqslant \sum_{r=1}^{n}p(p-1)\cdots(p-r+1)\sum_{P(\boldsymbol{\alpha},r)}\frac{\boldsymbol{\alpha}!}{\prod_{j=1}^{n}k_j!\boldsymbol{\beta}_j!}\bigg\|u^{p-r}(\cdot,\mathbf{y})\prod_{j=1}^{n}\Big(\partial_{\mathbf{y}}^{\boldsymbol{\beta}_j}u(\cdot,\mathbf{y})\Big)^{k_j}\bigg\|_{W^{1,1}(D)}\\ &\lesssim p\bigg(\frac{\boldsymbol{\gamma}}{\log 2}\bigg)^{\boldsymbol{\alpha}}\bigg(\frac{a_{\min}(\mathbf{y})}{a_{\min}(\mathbf{y})}\bigg)^p\|u(\cdot,\mathbf{y})\|_{W^{1,p}(D)}^p\sum_{r=1}^{n}\frac{p!}{(p-r)!}\sum_{P(\boldsymbol{\alpha},r)}\frac{\boldsymbol{\alpha}!}{\prod_{j=1}^{n}k_j!\boldsymbol{\beta}_j!}\prod_{j=1}^{n}\big(|\boldsymbol{\beta}_j|!\big)^{k_j}.\end{aligned}$$

From [9], we know that

$$\sum_{P(\boldsymbol{\alpha},r)}\frac{\boldsymbol{\alpha}!}{\prod_{j=1}^{n}k_j!\boldsymbol{\beta}_j!} = S_{n,r},$$

where $S_{n,r}$ are the Stirling numbers of the second kind, cf. [1]. Moreover, since $\prod_{j=1}^{n}\big(|\boldsymbol{\beta}_j|!\big)^{k_j}\leqslant|\boldsymbol{\alpha}|!$, we can further estimate

$$\sum_{r=1}^{n}\frac{p!}{(p-r)!}\sum_{P(\boldsymbol{\alpha},r)}\frac{\boldsymbol{\alpha}!}{\prod_{j=1}^{n}k_j!\boldsymbol{\beta}_j!}\prod_{j=1}^{n}\big(|\boldsymbol{\beta}_j|!\big)^{k_j}\leqslant|\boldsymbol{\alpha}|!\sum_{r=1}^{n}\frac{p!}{(p-r)!}S_{n,r} = |\boldsymbol{\alpha}|!p^n.$$

The last inequality follows by the theory of generating functions for the Stirling numbers of the second kind, see e.g. [1]. This finally completes the proof. ∎

4. quasi-Monte Carlo Quadrature for the Stochastic Variable

In this section, we discuss the use of quasi-Monte Carlo quadrature rules for the integral

$$\mathbf{I}v := \int_{(0,1)^m} v(\mathbf{z})\,\mathrm{d}\mathbf{z}.$$

These quadrature rules are classically of the form

$$\mathbf{Q}v := \frac{1}{N}\sum_{i=1}^{N} v(\boldsymbol{\xi}^i), \tag{25}$$

where N denotes the number of samples and $\boldsymbol{\xi}^i \in [0,1]^m$ denotes a single sample point.

The error estimation of the quasi-Monte Carlo method is usually performed for functions $f : [0,1]^m \to \mathbb{R}$ of *bounded variation in the sense of Hardy and Krause*, i.e.

$$V_{HK}(f) := \sum_{\|\boldsymbol{\alpha}\|_\infty = 1} V^{(|\boldsymbol{\alpha}|)}\big(f(\mathbf{y}_{\boldsymbol{\alpha}}, \mathbf{1})\big),$$

where $V^{(m)}(f)$ is the *variation of f in the sense of Vitali* on $[0,1]^m$. For a given vector $\mathbf{y} \in \mathbb{R}^m$, we denote by $\mathbf{y}_{\boldsymbol{\alpha}} \in \mathbb{R}^{|\boldsymbol{\alpha}|}$ the compressed vector containing the components y_k from $\mathbf{y}$ with $\alpha_k = 1$. Additionally, for $\mathbf{z} \in \mathbb{R}^m$ we write $(\mathbf{y}_{\boldsymbol{\alpha}}, \mathbf{z}) \in \mathbb{R}^m$ for the vector whose k-th component is given by y_k if $\alpha_k = 1$ and is given by z_k if $\alpha_k = 0$. For $\mathbf{z} = \mathbf{1}$, this vector is contained in the $|\boldsymbol{\alpha}|$-dimensional face $\{\mathbf{y} \in [0,1]^m : y_j = 1 \text{ for } \alpha_j = 0\}$, see [32]. Thus, $f(\mathbf{y}_{\boldsymbol{\alpha}}, \mathbf{1})$ corresponds to the restriction of f to this $|\boldsymbol{\alpha}|$-dimensional face. In addition, the norm $\|\boldsymbol{\alpha}\|_\infty$ of the multi-index $\boldsymbol{\alpha} \in \mathbb{N}^m$ is defined as usually subject to $\|\boldsymbol{\alpha}\|_\infty = \max_{k=1,\ldots,m} |\alpha_k|$.

The variation in the sense of Vitali has a simple expression if the function f has continuous partial derivatives. Then, it holds

$$V^{(m)}(f) = \int_{[0,1]^m} \left|\frac{\partial^m f}{\partial y_1 \cdots \partial y_m}(\mathbf{y})\right| \mathrm{d}\mathbf{y}.$$

Hence, the variation in the sense of Hardy and Krause can be written as

$$V_{HK}(f) = \sum_{\|\boldsymbol{\alpha}\|_\infty = 1} \int_{[0,1]^{|\boldsymbol{\alpha}|}} \big|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} f(\mathbf{y}_{\boldsymbol{\alpha}}, \mathbf{1})\big|\,\mathrm{d}\mathbf{y}_{\boldsymbol{\alpha}}.$$

Then, the error of a quasi-Monte Carlo method over the unit cube $[0,1]^m$ is bounded by means of the *star discrepancy* of the set of sample points $\Xi_N = \{\boldsymbol{\xi}^1, \ldots, \boldsymbol{\xi}^N\} \subset [0,1]^m$,

$$\mathcal{D}_\infty^\star(\Xi_N) := \sup_{\mathbf{t} \in [0,1]^m} |\mathrm{discr}_{\Xi_N}(\mathbf{t})|,$$

where the *local discrepancy function* $\mathrm{discr}_{\Xi_N} : [0,1]^m \to \mathbb{R}$ is defined by

$$\mathrm{discr}_{\Xi_N}(\mathbf{t}) := \mathrm{Vol}\big([\mathbf{0}, \mathbf{t})\big) - \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}_{[\mathbf{0},\mathbf{t})}\big(\boldsymbol{\xi}^i\big).$$

Here, $\mathrm{Vol}\big([\mathbf{0}, \mathbf{t})\big)$ denotes the Lebesgue measure of the cuboid $[\mathbf{0}, \mathbf{t})$. In addition, we mean by

$$\mathbb{1}_B(\mathbf{y}) = \begin{cases} 1, & \text{if } \mathbf{y} \in B, \\ 0, & \text{else,} \end{cases}$$

the indicator function of the set $B \subset \mathbb{R}^m$.

More precisely, the quadrature error can be estimated for functions with bounded variation in the sense of Hardy and Krause by

$$|(\mathbf{I} - \mathbf{Q}_{\Xi_N})f| \leqslant \mathcal{D}_\infty^\star(\Xi_N) V_{HK}(f), \tag{26}$$

which is the *Koksma-Hlawka* inequality, cf. [32]. In case of certain, so-called *low discrepancy* point sequences, e.g. the Sobol sequence, Niederreiter sequence or Halton sequence, this discrepancy can typically be estimated to be of the order $\mathcal{O}\big(N^{-1}(\log N)^m\big)$, see e.g. [3, 16, 32].

The estimate (26) is derived from the *Zaremba-Hlawka* identity, see [25, 47],

$$(\mathbf{I}-\mathbf{Q}_{\Xi_N})f = \sum_{\|\boldsymbol{\alpha}\|_\infty=1}(-1)^{|\boldsymbol{\alpha}|}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} f(\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{1})\operatorname{discr}_{\Xi_N}(\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{1})\,\mathrm{d}\mathbf{y}_{\boldsymbol{\alpha}} \tag{27}$$

by Hölder's inequality for sums and for integrals. Whenever the integrand provides some anisotropic behaviour, which means that some of the dimensions are of greater importance for the integrand than other dimensions, one can introduce weights into (27) which yields

$$(\mathbf{I}-\mathbf{Q}_{\Xi_N})f = \sum_{\|\boldsymbol{\alpha}\|_\infty=1}(-1)^{|\boldsymbol{\alpha}|}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} f(\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{1})w_{\boldsymbol{\alpha}}^{-1/2}w_{\boldsymbol{\alpha}}^{1/2}\operatorname{discr}_{\Xi_N}(\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{1})\,\mathrm{d}\mathbf{y}_{\boldsymbol{\alpha}}.$$

The application of Hölder's inequality for the integral as well as for the sum in the above equation yields the generalized, weighted Koksma-Hlawka inequality, see [27],

$$\big|(\mathbf{I}-\mathbf{Q}_{\Xi_N})f\big| \leqslant \mathcal{D}_{\mathbf{w}}^{r,s}(\Xi_N)\|f\|_{\mathcal{W}_{\mathbf{w}}^{r',s'}}, \tag{28}$$

with dual exponents r, r' and s, s' respectively. The weighted discrepancy $\mathcal{D}_{\mathbf{w}}^{r,s}(\Xi_N)$ is defined by

$$\mathcal{D}_{\mathbf{w}}^{r,s}(\Xi_N) := \left(\sum_{\|\boldsymbol{\alpha}\|_\infty=1}\|w_{\boldsymbol{\alpha}}^{1/2}\operatorname{discr}_{\Xi_N}(\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{1})\|_{L^r([0,1]^m)}^s\right)^{\frac{1}{s}} \tag{29}$$

and the norm $\|\cdot\|_{\mathcal{W}_{\mathbf{w}}^{r',s'}}$ by

$$\|f\|_{\mathcal{W}_{\mathbf{w}}^{r',s'}} := \left(\sum_{\|\boldsymbol{\alpha}\|_\infty=1}\|w_{\boldsymbol{\alpha}}^{-1/2}\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} f(\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{1})\|_{L^{r'}([0,1]^m)}^{s'}\right)^{\frac{1}{s'}}. \tag{30}$$

The modifications for the cases $r, s \in \{1,\infty\}$ are defined as usual. The norm here defines a Banach space $\mathcal{W}_{\mathbf{w}}^{r',s'}$. Especially, the integration error in this Banach space is then bounded by means of a weighted discrepancy.

Remark 4.1. *The estimation of the discrepancy or the weighted discrepancy of a set $\Xi \subset [0,1]^m$, especially for high dimensions m, has been the topic of many publications in the past fifteen years. The aim is to avoid the factor $(\log N)^m$ in the estimation of the discrepancy which grows exponentially in the dimension m. In certain cases, it is possible to construct point sequences such that the exponential dependence on the dimensionality can be removed, cf.* [34, 35, 36, 44, 46]. *Then, the integration problem in the corresponding Banach space is said to be tractable.*

The identity (27) is derived for functions $f : [0,1]^m \to \mathbb{R}$. In our applications, we would like to use quasi-Monte Carlo quadrature methods for the approximation of the solution's moments which appear as Bochner integrals. Thus, we consider functions $v \in C^1\big([0,1]^m; W^{1,q}(D)\big)$ for $q \in \mathbb{N}\backslash\{0\}$. It holds that the function $v(\mathbf{x},\cdot)$ is for almost all $\mathbf{x} \in D$ a continuous differentiable function $v(\mathbf{x},\cdot)\colon [0,1]^m \to \mathbb{R}$. Thus, we can apply for almost all $\mathbf{x} \in D$ the Zaremba-Hlawka identity (27) which yields

$$\big\|(\mathbf{I}-\mathbf{Q}_{\Xi_N})v\big\|_{W^{1,q}(D)} = \left\|\sum_{\|\boldsymbol{\alpha}\|_\infty=1}(-1)^{|\boldsymbol{\alpha}|}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} v(\cdot,\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{1})\operatorname{discr}_{\Xi_N}(\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{1})\,\mathrm{d}\mathbf{y}_{\boldsymbol{\alpha}}\right\|_{W^{1,q}(D)}. \tag{31}$$

Then, we obtain from the Bochner-inequality that

$$\begin{aligned}\big\|(\mathbf{I}-\mathbf{Q}_{\Xi_N})v\big\|_{W^{1,q}(D)} &\leqslant \sum_{\|\boldsymbol{\alpha}\|_\infty=1}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\big\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} v(\cdot,\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{1})\operatorname{discr}_{\Xi_N}(\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{1})\big\|_{W^{1,q}(D)}\,\mathrm{d}\mathbf{y}_{\boldsymbol{\alpha}}\\ &\leqslant \left(\sum_{\|\boldsymbol{\alpha}\|_\infty=1}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\big\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} v(\cdot,\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{1})\big\|_{W^{1,q}(D)}\,\mathrm{d}\mathbf{y}_{\boldsymbol{\alpha}}\right)\mathcal{D}_\infty^\star(\Xi_N).\end{aligned}$$

This is the analogue to the Koksma-Hlawka inequality for the evaluation of Bochner-integrals in $W^{1,q}(D)$. Of course, one can analogously obtain a generalized and weighted version of this inequality.

In the following, we assume that the sequence of integration points is given by the *Halton sequence.*

Definition 4.2. *Let $b_1, \dots, b_m$ denote the first m prime numbers. The m-dimensional* Halton sequence *is given by*

$$\boldsymbol{\xi}^i = [h_{b_1}(i), \dots, h_{b_m}(i)]^\intercal, \quad i = 0, 1, 2, \dots,$$

where $h_{b_j}(i)$ denotes the i-th element of the van der Corput sequence *with respect to b_j. That is, if $i = \cdots c_3 c_2 c_1$ in radix b_j, then $h_{b_j}(i) = 0.c_1 c_2 c_3 \cdots$ in radix b_j.*

We show that the convergence rate of the quasi-Monte Carlo quadrature based on this sequence for the determination of the moments of the solution u to (12) depends only linearly on the dimensionality m under certain decay properties of the sequence $\{\gamma_k\}_k$. The proof is essentially based on the ideas in [37].

To obtain a quasi-Monte Carlo method for the integration domain $\mathbb{R}^m$, we map the quadrature points to $\mathbb{R}^m$ by the inverse distribution function. This is equivalent to the transformation of the considered integrals to the unit cube. To that end, we define the cumulative normal distribution

$$\Phi \colon \mathbb{R} \to (0,1), \quad \text{with} \quad \Phi(y) := \int_{-\infty}^{y} \rho(y') \,\mathrm{d}y'$$

and its inverse

$$\Phi^{-1} \colon (0,1) \to \mathbb{R}.$$

Then, for a function $f \in L^1_\rho(\mathbb{R})$, it is well known that

$$\int_{\mathbb{R}} f(y)\rho(y) \,\mathrm{d}y = \int_0^1 f\big(\Phi^{-1}(z)\big) \,\mathrm{d}z$$

due to the substitution $z = \Phi(y)$. Especially, we have $f \circ \Phi^{-1} \in L^1\big((0,1)\big)$. By defining $\boldsymbol{\Phi}(\mathbf{y}) := [\Phi(y_1), \dots, \Phi(y_m)]^\intercal$, we may extend the above integral transform to the multivariate case, i.e. $f \in L^1_{\boldsymbol{\rho}}(\mathbb{R}^m)$ and

$$\int_{\mathbb{R}^m} f(\mathbf{y})\boldsymbol{\rho}(\mathbf{y}) \,\mathrm{d}\mathbf{y} = \int_{(0,1)^m} f\big(\boldsymbol{\Phi}^{-1}(\mathbf{z})\big) \,\mathrm{d}\mathbf{z}.$$

Although we have $f \circ \boldsymbol{\Phi}^{-1} \in L^1\big((0,1)^m\big)$, the integrand might be unbounded in a neighbourhood of the hypercube's boundary in our application since the diffusion coefficient may tend to zero. This implies that the variation in the sense of Hardy and Krause might be unbounded, too. As a consequence, the Koksma-Hlawka inequality is not applicable. The idea of [37] is now to consider subsets K_N such that the first N points $\boldsymbol{\xi}^1, \dots, \boldsymbol{\xi}^N$ of the Halton sequence are included in K_N. Due to the definition of the Halton sequence, this holds for the cuboid

$$K_N := \bigtimes_{k=1}^{m} [(b_k N)^{-1}, 1 - (b_k N)^{-1}].$$

Obviously, for the solution u to (9), it holds for almost every $\mathbf{x} \in D$

$$\operatorname*{ess\,sup}_{\mathbf{z} \in K_N} u\big(\mathbf{x}, \boldsymbol{\Phi}^{-1}(\mathbf{z})\big) < \infty \quad \text{for all } N \in \mathbb{N}.$$

Let now $\hat{u}(\mathbf{x},\mathbf{z}) := u\big(\mathbf{x}, \boldsymbol{\Phi}^{-1}(\mathbf{z})\big)$. For $\mathbf{z} \in (0,1)^m \backslash K_N$ and almost every $\mathbf{x} \in D$, we replace $\hat{u}$ by its *low-variation extension*

$$\hat{u}_{\text{ext}}(\mathbf{x},\mathbf{z}) := \hat{u}(\mathbf{c}) + \sum_{\|\boldsymbol{\alpha}\|_\infty = 1} \int_{[\mathbf{c}_{\boldsymbol{\alpha}}, \mathbf{z}_{\boldsymbol{\alpha}}]} \mathbb{1}_{(\mathbf{y}_{\boldsymbol{\alpha}}, \mathbf{c}) \in K_N} \partial^{\boldsymbol{\alpha}}_{\mathbf{y}} \hat{u}(\mathbf{x}, \mathbf{y}_{\boldsymbol{\alpha}}, \mathbf{c}) \,\mathrm{d}\mathbf{y}_{\boldsymbol{\alpha}}, \tag{32}$$

where

$$\mathbb{1}_{(\mathbf{y}_{\boldsymbol{\alpha}}, \mathbf{c}) \in K_N} := \begin{cases} 1, & \text{if } (\mathbf{y}_{\boldsymbol{\alpha}}, \mathbf{c}) \in K_N, \\ 0, & \text{else,} \end{cases}$$

denotes the indicator function of the set K_N.

For a given *anchor point* $\mathbf{c} \in K_N$, the extension coincides by definition (32) with the function $\hat{u}$ on K_N, i.e. $\hat{u}_{\text{ext}}(\mathbf{x},\mathbf{z}) = \hat{u}(\mathbf{x},\mathbf{z})$ for all $\mathbf{z} \in K_N$. We are now ready to prove our main result.

Theorem 4.3. *The quasi-Monte Carlo quadrature using Halton points for approximating the expectation of the solution u to* (9) *provides a convergence rate which depends only linearly on the dimensionality m if the sequence $\{\gamma_k\}_k$ satisfies the decay property $\gamma_k \lesssim k^{-4-2\eta}$ for an arbitrary $\eta > 0$. More precisely, there exists for each $\delta > 0$ a sequence $\{\delta_k\}_{k>0} \in \ell^1(\mathbb{N})$ with $\delta_k \eqsim k^{-1-\eta}$ and a $\tilde{\delta} > 0$ with $\tilde{\delta} + \sum_{k=1}^{\infty} \delta_k < \delta$ such that the error of the quasi-Monte Carlo quadrature with N Halton points satisfies*

$$\begin{aligned} \big\|(\mathbf{I} - \mathbf{Q}_{\Xi_N})\hat{u}\big\|_{H^1(D)} &\lesssim \|f\|_{L^2(D)}\big(mN^{-1+\|\boldsymbol{\delta}\|_\infty} + N^{-1+\tilde{\delta}+|\boldsymbol{\delta}|}\big) \\ &\leqslant \|f\|_{L^2(D)}(m+1)N^{-1+\delta}. \end{aligned} \tag{33}$$

Herein, we denote by $\boldsymbol{\delta} = [\delta_1, \ldots, \delta_m]^\intercal$ the vector which contains the first m entries of the series $\{\delta_k\}_{k>0}$. The constant hidden in the above inequality depends on $\{\delta_k\}_{k>0}$, on $\tilde{\delta}$ and on δ.

The proof of this theorem is performed by splitting up the error of integration into three parts. Namely, with respect to the extension $\hat{u}_{\text{ext}}$, we write

$$\big\|(\mathbf{I} - \mathbf{Q}_{\Xi_N})\hat{u}\big\|_{H^1(D)} \leqslant \big\|\mathbf{I}(\hat{u} - \hat{u}_{\text{ext}})\big\|_{H^1(D)} + \big\|\mathbf{Q}_{\Xi_N}(\hat{u} - \hat{u}_{\text{ext}})\big\|_{H^1(D)} + \big\|(\mathbf{I} - \mathbf{Q}_{\Xi_N})\hat{u}_{\text{ext}}\big\|_{H^1(D)}. \tag{34}$$

Due to $\hat{u}\big|_{K_N} = \hat{u}_{\text{ext}}\big|_{K_N}$, the second term on the right-hand side vanishes in this inequality. The first term on the right-hand side of (34), which corresponds to the truncation error of the quasi-Monte Carlo quadrature based on the Halton sequence, is estimated by Lemma 4.4. The third term on the right-hand side of (34), which reflects the integration error inside K_N, is estimated in Lemma 4.7.

Lemma 4.4. *Let the conditions of Theorem 4.3 hold and let $\hat{u}_{\text{ext}}$ be defined according to* (32). *Then, it holds*

$$\big\|\mathbf{I}(\hat{u} - \hat{u}_{\text{ext}})\big\|_{H^1(D)} \lesssim \|f\|_{L^2(D)} N^{-1+\|\boldsymbol{\delta}\|_\infty} m. \tag{35}$$

Proof. We organize the proof in four steps.
(*i.*) On the one hand, from [10], we know that

$$\Phi^{-1}(z) < \sqrt{-\log\big(2\pi(1-z)^2(1-\log(2\pi(1-z)^2))\big)} \quad \text{for all } z \in [0.9, 1].$$

Furthermore, we have from [38] that

$$\Phi^{-1}(z) \leqslant \sqrt{-2\log(1-z)} - \frac{2.30753 + 0.27061\sqrt{-2\log(1-z)}}{1 + 0.99229\sqrt{-2\log(1-z)} - 0.08962\log(1-z)} + 0.003$$

for all $z \in [0.5, 1]$. These inequalities imply

$$\Phi^{-1}(z) \leqslant \sqrt{-2\log(1-z)} \quad \text{for all } z \in [0.5, 1].$$

Due to the symmetry of the distribution, this shows that

$$|\Phi^{-1}(z)| \leqslant \sqrt{-2\log(\min\{z, 1-z\})} \quad \text{for all } z \in [0, 1].$$

The derivative of the distribution function is the Gaussian density function. Hence, the derivative of its inverse can easily be determined. Therefore, we derive

$$\frac{\mathrm{d}}{\mathrm{d}z}\Phi^{-1}(z) = \sqrt{2\pi}\exp\left(\frac{\Phi^{-1}(z)^2}{2}\right) \leqslant \sqrt{2\pi}\min\{z, 1-z\}^{-1},$$

which implies the estimate

$$\left|\prod_{k=1}^{m}\left(\frac{\mathrm{d}}{\mathrm{d}z_k}\Phi^{-1}(z_k)\right)^{\alpha_k}\right| \leqslant \prod_{k=1}^{m}\big(\sqrt{2\pi}\min\{z_k, 1-z_k\}^{-1}\big)^{\alpha_k}$$

for all non-negative integers $\alpha_k \geqslant 0$.
(*ii.*) On the other hand, one verifies

$$\exp\big(\gamma_k|\Phi^{-1}(z)|\big) \leqslant C(\delta_k, \gamma_k)\min\{z, 1-z\}^{-\delta_k} \quad \text{for all } \delta_k > 0,$$

with the constant

$$C(\delta_k,\gamma_k) = \begin{cases} \exp\left(\frac{\gamma_k^2}{2\delta_k}\right), & \text{if } \delta_k \leqslant \frac{\gamma_k}{\sqrt{2\log 2}},\\ \frac{\exp(\sqrt{2\log 2}\gamma_k)}{\exp(\delta_k \log 2)}, & \text{else.}\end{cases}$$

Hence, we find by the definition of $a_{\max}$ and $a_{\min}$ that

$$\sqrt{\frac{a_{\max}\big(\boldsymbol{\Phi}^{-1}(\mathbf{z})\big)}{a_{\min}\big(\boldsymbol{\Phi}^{-1}(\mathbf{z})\big)^3}} \leqslant \exp\left(\sum_{k=1}^m 2\gamma_k\big|\Phi^{-1}(z_k)\big|\right) \leqslant \prod_{k=1}^m \big(C(\delta_k,2\gamma_k)\min\{z_k,1-z_k\}^{-\delta_k}\big).$$

Consequently, with Lemma 3.1 and the stability estimate (13), for any multi-index $\boldsymbol{\alpha}$, we deduce

$$\begin{aligned}\big\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u\big(\cdot,\boldsymbol{\Phi}^{-1}(\mathbf{z})\big)\big\|_{H^1(D)} &\leqslant |\boldsymbol{\alpha}|!\left(\frac{\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}}\sqrt{\frac{a_{\max}\big(\boldsymbol{\Phi}^{-1}(\mathbf{z})\big)}{a_{\min}\big(\boldsymbol{\Phi}^{-1}(\mathbf{z})\big)}}\big\|u\big(\cdot,\boldsymbol{\Phi}^{-1}(\mathbf{z})\big)\big\|_{H^1(D)}\\ &\leqslant |\boldsymbol{\alpha}|!\left(\frac{\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}}\sqrt{\frac{a_{\max}\big(\boldsymbol{\Phi}^{-1}(\mathbf{z})\big)}{a_{\min}\big(\boldsymbol{\Phi}^{-1}(\mathbf{z})\big)^3}}\|f\|_{L^2(D)}\\ &\leqslant \|f\|_{L^2(D)}|\boldsymbol{\alpha}|!\left(\frac{\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}}\prod_{k=1}^m\Big(C(\delta_k,2\gamma_k)\min\{z_k,1-z_k\}^{-\delta_k}\Big).\end{aligned}$$

(*iii.*) For an arbitrary multi-index $\boldsymbol{\alpha}$, it holds for all $\mathbf{z}\in(0,1)^m$ that

$$\begin{aligned}\big\|\partial_{\mathbf{z}}^{\boldsymbol{\alpha}}\hat{u}(\cdot,\mathbf{z})\big\|_{H^1(D)} &= \big\|\partial_{\mathbf{z}}^{\boldsymbol{\alpha}} u\big(\cdot,\boldsymbol{\Phi}^{-1}(\mathbf{z})\big)\big\|_{H^1(D)}\\ &= \left\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u\big(\cdot,\boldsymbol{\Phi}^{-1}(\mathbf{z})\big)\prod_{k=1}^m\left(\frac{\mathrm{d}}{\mathrm{d}z_k}\Phi^{-1}(z_k)\right)^{\alpha_k}\right\|_{H^1(D)}\\ &= \left|\prod_{k=1}^m\left(\frac{\mathrm{d}}{\mathrm{d}z_k}\Phi^{-1}(z_k)\right)^{\alpha_k}\right|\big\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u\big(\cdot,\boldsymbol{\Phi}^{-1}(\mathbf{z})\big)\big\|_{H^1(D)}.\end{aligned}\tag{36}$$

From now on, we choose the anchor point $\mathbf{c}=\mathbf{1/2}$ and define

$$\tilde{C} := \frac{\sqrt{2\pi}\max_{k\in\mathbb{N}} C(\delta_k,2\gamma_k)}{\log 2}.\tag{37}$$

Note that $\tilde{C}<\infty$ since there is a $k_0\in\mathbb{N}$ such that $C(\delta_k,2\gamma_k)\leqslant 1$ for all $k\geqslant k_0$ under the decay assumptions on the sequences $\{\delta_k\}_k$ and $\{\gamma_k\}_k$. Due to $\Phi^{-1}(1/2)=0$, we easily get from item (*ii.*), Lemma 3.1 and (13) that

$$\big\|\partial_{\mathbf{y}}^{\boldsymbol{\alpha}} u\big(\cdot,\boldsymbol{\Phi}^{-1}(\mathbf{z}_{\boldsymbol{\alpha}},\mathbf{c})\big)\big\|_{H^1(D)} \lesssim \|f\|_{L^2(D)}|\boldsymbol{\alpha}|!\left(\frac{\boldsymbol{\gamma}}{\log 2}\right)^{\boldsymbol{\alpha}}\prod_{k=1}^m\Big(C(\delta_k,2\gamma_k)\min\{z_k,1-z_k\}^{-\delta_k}\Big)^{\alpha_k}\tag{38}$$

holds for all $\boldsymbol{\alpha}$ with $\|\boldsymbol{\alpha}\|_\infty=1$. Thus, by combining (36) with item (*i.*) and inequality (38), we arrive at the estimate

$$\big\|\partial_{\mathbf{z}}^{\boldsymbol{\alpha}}\hat{u}(\cdot,\mathbf{z}_{\boldsymbol{\alpha}},\mathbf{c})\big\|_{H^1(D)} \lesssim |\boldsymbol{\alpha}|!\|f\|_{L^2(D)}\prod_{k=1}^m\Big(\gamma_k\tilde{C}\min\{z_k,1-z_k\}^{-1-\delta_k}\Big)^{\alpha_k}.\tag{39}$$

From (32), we infer the identity

$$\hat{u}(\cdot,\mathbf{z})-\hat{u}_{\mathrm{ext}}(\cdot,\mathbf{z}) = \sum_{\|\boldsymbol{\alpha}\|_\infty=1}\int_{[\mathbf{c}_{\boldsymbol{\alpha}},\mathbf{z}_{\boldsymbol{\alpha}}]}\mathbb{1}_{(\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{c})\notin K_N}\partial^{\boldsymbol{\alpha}}\hat{u}(\cdot,\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{c})\,\mathrm{d}\mathbf{y}_{\boldsymbol{\alpha}}.$$

This, together with the estimate (39) on the derivates of $\hat{u}$ yields for every $\mathbf{z} \notin K_N$, cf. [37],

$$\begin{aligned}
&\|\hat{u}(\cdot,\mathbf{z}) - \hat{u}_{\text{ext}}(\cdot,\mathbf{z})\|_{H^1(D)} \\
&\quad \leqslant \sum_{\|\boldsymbol{\alpha}\|_\infty=1} \int_{[\mathbf{c}_{\boldsymbol{\alpha}},\mathbf{z}_{\boldsymbol{\alpha}}]} \mathbb{1}_{(\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{c})\notin K_N} \big\|\partial^{\boldsymbol{\alpha}} \hat{u}(\cdot,\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{c})\big\|_{H^1(D)} \,\mathrm{d}\mathbf{y}_{\boldsymbol{\alpha}} \\
&\quad \lesssim \|f\|_{L^2(D)} \sum_{\|\boldsymbol{\alpha}\|_\infty=1} |\boldsymbol{\alpha}|! \prod_{k=1}^m \big(\gamma_k \tilde{C}\big)^{\alpha_k} \int_{[\mathbf{c}_{\boldsymbol{\alpha}},\mathbf{z}_{\boldsymbol{\alpha}}]} \mathbb{1}_{(\mathbf{y}_{\boldsymbol{\alpha}},\mathbf{c})\notin K_N} \prod_{k=1}^m \Big(\min\{y_k, 1-y_k\}^{-1-\delta_k}\Big)^{\alpha_k} \mathrm{d}\mathbf{y}_{\boldsymbol{\alpha}} \\
&\quad \leqslant \|f\|_{L^2(D)} \sum_{\|\boldsymbol{\alpha}\|_\infty=1} |\boldsymbol{\alpha}|! \prod_{k=1}^m \bigg(\gamma_k \tilde{C} \int_{\min\{z_k,1-z_k\}}^{1/2} y_k^{-1-\delta_k} \,\mathrm{d}y_k\bigg)^{\alpha_k}.
\end{aligned}$$

Herein, the integral can simply be bounded via its lower limit according to

$$\begin{aligned}
\|\hat{u}(\cdot,\mathbf{z}) - \hat{u}_{\text{ext}}(\cdot,\mathbf{z})\|_{H^1(D)} &\lesssim \|f\|_{L^2(D)} \sum_{\|\boldsymbol{\alpha}\|_\infty=1} |\boldsymbol{\alpha}|! \prod_{k=1}^m \big(\gamma_k \tilde{C} \min\{z_k, 1-z_k\}^{-\delta_k}\big)^{\alpha_k} \\
&\leqslant \|f\|_{L^2(D)} \sum_{\|\boldsymbol{\alpha}\|_\infty=1} \prod_{k=1}^m \big(k\gamma_k \tilde{C} \min\{z_k, 1-z_k\}^{-\delta_k}\big)^{\alpha_k} \\
&= \|f\|_{L^2(D)} \Bigg(\prod_{k=1}^m \bigg(1 + \frac{\min\{z_k, 1-z_k\}^{-\delta_k} k\gamma_k \tilde{C}}{\delta_k}\bigg) - 1\Bigg) \\
&\leqslant \|f\|_{L^2(D)} \prod_{k=1}^m \bigg(1 + \frac{k\gamma_k \tilde{C}}{\delta_k}\bigg) \min\{z_k, 1-z_k\}^{-\delta_k}.
\end{aligned} \tag{40}$$

Now, due to Bochner's inequality and due to the fact that $\hat{u}$ coincides with $\hat{u}_{\text{ext}}$ in K_N, it follows

$$\big\|\mathbf{I}(\hat{u} - \hat{u}_{\text{ext}})\big\|_{H^1(D)} \leqslant \int_{(0,1)^m} \big\|\hat{u}(\cdot,\mathbf{z}) - \hat{u}_{\text{ext}}(\cdot,\mathbf{z})\big\|_{H^1(D)} \,\mathrm{d}\mathbf{z} = \int_{(0,1)^m \setminus K_N} \big\|\hat{u}(\cdot,\mathbf{z}) - \hat{u}_{\text{ext}}(\cdot,\mathbf{z})\big\|_{H^1(D)} \,\mathrm{d}\mathbf{z}.$$

With the estimate (40), this implies

$$\begin{aligned}
\big\|\mathbf{I}(\hat{u} - \hat{u}_{\text{ext}})\big\|_{H^1(D)} &\lesssim \|f\|_{L^2(D)} \int_{(0,1)^m \setminus K^N} \prod_{k=1}^m \min\{z_k, 1-z_k\}^{-\delta_k} \,\mathrm{d}\mathbf{z} \prod_{k=1}^m \bigg(1 + \frac{k\gamma_k \tilde{C}}{\delta_k}\bigg) \\
&\leqslant \|f\|_{L^2(D)} 2^m \sum_{j=1}^m \int_0^{(b_j N)^{-1}} z_j^{-\delta_j} \,\mathrm{d}z_j \prod_{i=1, i\neq j}^m \int_0^{1/2} z_i^{-\delta_i} \,\mathrm{d}z_i \prod_{k=1}^m \bigg(1 + \frac{k\gamma_k \tilde{C}}{\delta_k}\bigg) \\
&\leqslant \|f\|_{L^2(D)} 2^m \sum_{j=1}^m (b_j N)^{\delta_j - 1} 2^{-m+1} 2^{|\boldsymbol{\delta}|} \prod_{k=1}^m \bigg[\bigg(1 + \frac{k\gamma_k \tilde{C}}{\delta_k}\bigg)\bigg(\frac{1}{1-\delta_k}\bigg)\bigg] \\
&\lesssim \|f\|_{L^2(D)} N^{\|\boldsymbol{\delta}\|_\infty - 1} m \prod_{k=1}^m \bigg[\bigg(1 + \frac{k\gamma_k \tilde{C}}{\delta_k}\bigg)\bigg(\frac{1}{1-\delta_k}\bigg) 2^{\delta_k}\bigg].
\end{aligned}$$

(*iv.*) It remains to prove that the appearing constants are bounded independently of the dimension m. Therefore, it is now sufficient to show

$$\prod_{k=1}^\infty \bigg[\bigg(1 + \frac{k\gamma_k \tilde{C}}{\delta_k}\bigg)\bigg(\frac{1}{1-\delta_k}\bigg) 2^{\delta_k}\bigg] < \infty. \tag{41}$$

Since we may choose $\delta_k > 0$ arbitrarily, we can assume that the sequence $\{\delta_k\}$ satisfies the conditions of Theorem 4.3. Then, it holds

$$\prod_{k=1}^\infty 2^{\delta_k} = 2^{\sum_{k=1}^\infty \delta_k} \leqslant 2^\delta \quad \text{and} \quad \prod_{k=1}^\infty \frac{1}{1-\delta_k} = \exp\bigg(-\sum_{k=1}^\infty \log(1-\delta_k)\bigg). \tag{42}$$

We make use of the fact that the Taylor expansion of the logarithm $\log(x)$ at $x = 1$ is given by

$$\log(1-h) = -\sum_{k=1}^\infty \frac{h^k}{k} = -h - \mathcal{O}(h^2), \quad h > 0.$$

By inserting this into the equation on the right of (42), we obtain

$$\prod_{k=1}^{\infty}\frac{1}{1-\delta_k}\leqslant\exp\Bigg(\sum_{k=1}^{\infty}\big(\delta_k+\mathcal{O}(\delta_k^2)\big)\Bigg)\lesssim\exp(\delta+c\delta^2)$$

for some $c>0$. Since the sequence $\{\gamma_k\}_k$ decays asymptotically faster than $k^{-4-2\eta}$, we conclude for some $c>0$ that

$$\prod_{k=1}^{\infty}\left(1+\frac{\tilde{C}k\gamma_k}{\delta_k}\right)\leqslant\prod_{k=1}^{\infty}\big(1+ck^{-2-\eta}\big)<\infty. \tag{43}$$

This establishes estimate (41) and, thus, finally the assertion (35). □

Remark 4.5. *Notice that the last estimate of item (iii.) is quite rough. In fact, if we sum up* $\sum_{j=1}^{m} b_j^{\delta_j-1}$*, we end up rather with a factor* $\log(m)$ *than a factor* m *since* $b_k\eqsim k\log k$*. Moreover, for this lemma, the weaker decay condition* $\{\gamma_k\}_k\lesssim k^{-3-2\eta}$ *is sufficient. This can be easily seen from equation* (43) *and the definition of the constant* $\tilde{C}$*, see* (37)*. These are the only parts in the proof of Lemma 4.4 where the decay properties of* $\{\gamma_k\}_k$ *enter. Especially, they remain valid under the weaker assumption* $\{\gamma_k\}_k\lesssim k^{-3-2\eta}$*.*

Finally, we bound the third term in (34). In [23], the *centered discrepancy* is introduced to establish an estimate for the error of integration. In the sequel, we will also make use of the *extreme discrepancy.*

Definition 4.6. *The* pointwise centered discrepancy *function is defined for a given set of* N *sample points* $\Xi_N\subset[0,1]^m$ *as* $\operatorname{discr}_c(\Xi_N)\colon[0,1]^m\to\mathbb{R}$ *with*

$$\operatorname{discr}_c(\mathbf{z},\Xi_N):=(-1)^{\sum_{k=1}^{m}\mathbb{1}_{\{z_k>1/2\}}}\Bigg(\prod_{k=1}^{m}\big(-z_k+\mathbb{1}_{\{z_k>1/2\}}\big)-\frac{1}{N}\sum_{i=1}^{N}\prod_{k=1}^{m}\big(\mathbb{1}_{\{z_k>1/2\}}-\mathbb{1}_{\{z_k>\xi_k^i\}}\big)\Bigg).$$

Then, the L^∞-centered discrepancy *is given by*

$$\mathcal{D}_c(\Xi_N):=\sup_{\mathbf{z}\in[0,1]^m}\big|\operatorname{discr}_c(\mathbf{z},\Xi_N)\big|.$$

Furthermore, the extreme discrepancy *is defined by*

$$\mathcal{D}_{\mathrm{extr}}(\Xi_N)=\sup_{\mathbf{x},\mathbf{y}\in[0,1]^m}\left|\operatorname{Vol}\big([\mathbf{x},\mathbf{y})\big)-\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}_{[\mathbf{x},\mathbf{y})}(\boldsymbol{\xi}^i)\right|.$$

Obviously, the L^∞-centered discrepancy is bounded by the extreme discrepancy. In the following, it is convenient to introduce the projection $(\Xi)_{\boldsymbol{\alpha}}$ of Ξ as $(\Xi)_{\boldsymbol{\alpha}}:=\{\xi_{\boldsymbol{\alpha}},\xi\in\Xi\}$.

Lemma 4.7. *Let the conditions of Theorem 4.3 be satisfied and let* $\hat{u}_{\mathrm{ext}}$ *be defined by* (32)*. Then, it holds*

$$\big\|(\mathbf{I}-\mathbf{Q}_{\Xi_N})\hat{u}_{\mathrm{ext}}\big\|_{H^1(D)}\lesssim\|f\|_{L^2(D)}N^{-1+\tilde{\delta}+|\boldsymbol{\delta}|}. \tag{44}$$

Proof. In [24], the Zaremba-Hlawka identity is generalized to discrepancy functions anchored at an arbitrary point $\mathbf{c}$ in $[0,1]^m$. This identity reads for the anchor point $\mathbf{c}=\mathbf{1/2}$ and a differentiable function $f:\mathbb{R}^m\to\mathbb{R}$ as follows:

$$(\mathbf{I}-\mathbf{Q}_{\Xi_N})f=\sum_{\|\boldsymbol{\alpha}\|\leqslant1}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\partial_{\mathbf{z}}^{\boldsymbol{\alpha}}f(\mathbf{z}_{\boldsymbol{\alpha}},\mathbf{c})\operatorname{discr}_c(\mathbf{z}_{\boldsymbol{\alpha}},(\Xi_N)_{\boldsymbol{\alpha}})\,\mathrm{d}\mathbf{z}_{\boldsymbol{\alpha}}. \tag{45}$$

Hence, we arrive for almost all $\mathbf{x}\in D$ at the representation of the quadrature error

$$(\mathbf{I}-\mathbf{Q}_{\Xi_N})\hat{u}_{\mathrm{ext}}(\mathbf{x})=\sum_{\|\boldsymbol{\alpha}\|\leqslant1}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\partial_{\mathbf{z}}^{\boldsymbol{\alpha}}\hat{u}_{\mathrm{ext}}(\mathbf{x},\mathbf{z}_{\boldsymbol{\alpha}},\mathbf{c})\operatorname{discr}_c\big(\mathbf{z}_{\boldsymbol{\alpha}},(\Xi_N)_{\boldsymbol{\alpha}}\big)\,\mathrm{d}\mathbf{z}_{\boldsymbol{\alpha}}.$$

Now, we obtain in the same way as in (31) the error estimate

$$\big\|(\mathbf{I}-\mathbf{Q}_{\Xi_N})\hat{u}_{\mathrm{ext}}\big\|_{H^1(D)}\leqslant\sum_{\|\boldsymbol{\alpha}\|_\infty=1}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\big\|\partial_{\mathbf{z}}^{\boldsymbol{\alpha}}\hat{u}_{\mathrm{ext}}(\cdot,\mathbf{z}_{\boldsymbol{\alpha}},\mathbf{c})\big\|_{H^1(D)}\,\mathrm{d}\mathbf{z}_{\boldsymbol{\alpha}}\sup_{\mathbf{z}_{\boldsymbol{\alpha}}\in[0,1]^{|\boldsymbol{\alpha}|}}\operatorname{discr}_c\big(\mathbf{z}_{\boldsymbol{\alpha}},(\Xi_N)_{\boldsymbol{\alpha}}\big).$$

In order to prove (44) with a constant independent of m, we introduce weights $w_k \in (0,\infty)$ for $k = 1,\dots,m$ and define the corresponding product weights with respect to the multi-index $\boldsymbol{\alpha}$ as $w_{\boldsymbol{\alpha}} := \prod_{k=1}^m w_k^{\alpha_k}$. Later on, we will specify these weights by exploiting the decay properties of the occurring derivatives of the integrand. From the above inequality, we deduce

$$\begin{aligned}\big\|(\mathbf{I}-\mathbf{Q}_{\Xi_N})\hat{u}_{\text{ext}}\big\|_{H^1(D)} &\leqslant \sum_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{-1/2}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\big\|\partial_{\mathbf{z}}^{\boldsymbol{\alpha}}\hat{u}_{\text{ext}}(\cdot,\mathbf{z}_{\boldsymbol{\alpha}},\mathbf{c})\big\|_{H^1(D)}\,\mathrm{d}\mathbf{z}_{\boldsymbol{\alpha}} w_{\boldsymbol{\alpha}}^{1/2}\mathcal{D}_c\big((\Xi_N)_{\boldsymbol{\alpha}}\big)\\ &\leqslant \bigg\{\sup_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{-1/2}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\big\|\partial_{\mathbf{z}}^{\boldsymbol{\alpha}}\hat{u}_{\text{ext}}(\cdot,\mathbf{z}_{\boldsymbol{\alpha}},\mathbf{c})\big\|_{H^1(D)}\,\mathrm{d}\mathbf{z}_{\boldsymbol{\alpha}}\bigg\}\bigg\{\sum_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{1/2}\mathcal{D}_c\big((\Xi_N)_{\boldsymbol{\alpha}}\big)\bigg\}.\end{aligned}$$

This corresponds in the terminology of the beginning of this section to the weighted centered Koksma-Hlawka inequality with the choices $r=\infty$ and $s=1$, see (28). Due to the definition of $\hat{u}_{\text{ext}}$, cf. (32), the derivative $\partial_{\mathbf{z}}^{\boldsymbol{\alpha}}\hat{u}_{\text{ext}}(\cdot,\mathbf{z}_{\boldsymbol{\alpha}},\mathbf{c})$ vanishes in $[0,1]^{|\boldsymbol{\alpha}|}\backslash\big(K_N\big)_{\boldsymbol{\alpha}}$ and coincides with the derivative of $\hat{u}$ in $\big(K_N\big)_{\boldsymbol{\alpha}}$. Therefore, with $\tilde{C}$ defined as in (37), we can estimate

$$\begin{aligned}&\sup_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{-1/2}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\big\|\partial_{\mathbf{z}}^{\boldsymbol{\alpha}}\hat{u}_{\text{ext}}(\cdot,\mathbf{z}_{\boldsymbol{\alpha}},\mathbf{c})\big\|_{H^1(D)}\,\mathrm{d}\mathbf{z}_{\boldsymbol{\alpha}}\\ &\quad\lesssim \|f\|_{L^2(D)}\sup_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{-1/2}|\boldsymbol{\alpha}|!\int_{(K_N)_{\boldsymbol{\alpha}}}\prod_{\{k:\alpha_k=1\}}\big(\gamma_k\tilde{C}\min\{z_k,1-z_k\}^{-1-\delta_k}\big)\,\mathrm{d}\mathbf{z}_{\boldsymbol{\alpha}}\\ &\quad\leqslant \|f\|_{L^2(D)}\sup_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{-1/2}2^{|\boldsymbol{\alpha}|}\prod_{\{k:\alpha_k=1\}}\bigg(k\gamma_k\tilde{C}\int_{(b_kN)^{-1}}^{1/2} z_k^{(-1-\delta_k)}\,\mathrm{d}z_k\bigg)\\ &\quad\leqslant \|f\|_{L^2(D)}\sup_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{-1/2}\prod_{k=1}^m\bigg(\frac{2k\gamma_k\tilde{C}}{\delta_k}\bigg)^{\alpha_k}(b_kN)^{\alpha_k\delta_k}.\end{aligned}$$

The choice of the weights

$$w_k = \frac{8\pi C(\delta_k,2\gamma_k)^2k^2\gamma_k^2}{\delta_k^2\log^2 2} \tag{46}$$

yields

$$w_{\boldsymbol{\alpha}}^{1/2} = \prod_{k=1}^m\bigg(\frac{2k\gamma_k\tilde{C}}{\delta_k}\bigg)^{\alpha_k}.$$

Therefore, we obtain

$$\sup_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{-1/2}\prod_{k=1}^m\bigg(\frac{2k\gamma_k\tilde{C}}{\delta_k}\bigg)^{\alpha_k}(b_kN)^{\alpha_k\delta_k}\leqslant N^{|\boldsymbol{\delta}|}\prod_{k=1}^m b_k^{\alpha_k\delta_k}.$$

Now, the prime number theorem, see e.g. [40], implies that $b_k < 2k\log(k+2)$. Hence, we deduce

$$\prod_{k=1}^\infty b_k^{\delta_k} = \exp\bigg(\sum_{k=1}^\infty \delta_k\log b_k\bigg)\lesssim \exp\bigg(\sum_{k=1}^\infty k^{-1-\eta}\log\big(2k\log(k+2)\big)\bigg)<\infty.$$

From this, we conclude that

$$\sup_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{-1/2}\int_{[0,1]^{|\boldsymbol{\alpha}|}}\big\|\partial_{\mathbf{z}}^{\boldsymbol{\alpha}}\hat{u}_{\text{ext}}(\cdot,\mathbf{z}_{\boldsymbol{\alpha}},\mathbf{c})\big\|_{H^1(D)}\,\mathrm{d}\mathbf{z}_{\boldsymbol{\alpha}}\lesssim N^{|\boldsymbol{\delta}|}\|f\|_{L^2(D)}.$$

In order to bound the weighted sum of the L^∞-centered discrepancies, we use the following result from [32]. It holds

$$\mathcal{D}_{\text{extr}}(\Xi_N)\leqslant 2^m\mathcal{D}^\star(\Xi_N).$$

Thus, we have

$$\sum_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{1/2}\mathcal{D}_c\big((\Xi_N)_{\boldsymbol{\alpha}}\big)\leqslant \sum_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{1/2}\mathcal{D}_{\text{extr}}\big((\Xi_N)_{\boldsymbol{\alpha}}\big)\leqslant \sum_{\|\boldsymbol{\alpha}\|_\infty=1} w_{\boldsymbol{\alpha}}^{1/2}2^{|\boldsymbol{\alpha}|}\mathcal{D}^\star\big((\Xi_N)_{\boldsymbol{\alpha}}\big).$$

Under the decay property

$$\sum_{k=1}^{\infty} \tilde{w}_k^{1/2} k \log k < \infty$$

of the weights $\tilde{w}_k := 4w_k$, it is shown in [45] that

$$\sum_{\|\boldsymbol{\alpha}\|_\infty = 1} \tilde{w}_{\boldsymbol{\alpha}}^{1/2} \mathcal{D}_\infty^\star\big((\Xi_N)_{\boldsymbol{\alpha}}\big) \lesssim N^{-1+\tilde{\delta}}$$

holds for all $\tilde{\delta} > 0$ with a constant which depends on $\tilde{\delta}$ but not on the dimensionality m. This condition is satisfied if the weights fulfill $\tilde{w}_k^{1/2} \lesssim k^{-2-\eta}$. Therefore, we get the following condition on the decay of γ_k:

$$\frac{4k\gamma_k \tilde{C}}{\delta_k} \lesssim k^{-2-\eta} \quad \Longrightarrow \quad \gamma_k \lesssim \frac{\delta_k}{4\tilde{C}} k^{-3-\eta} \sim k^{-4-2\eta}.$$

□

With the preceding two Lemmata at hand, we can establish the estimate (34). This completes the proof of Theorem 4.3.

Remark 4.8. *In this section, we have only shown approximation results of the quasi-Monte Carlo quadrature based on Halton points for the mean of the function* u*, i.e.* $\mathbb{E}_u$*. Note that if* $f \in L^p(D)$ *due to the regularity estimates proven in Section 3, the results in this section hold in complete analogy for the* p*-th moment* $\mathcal{M}^p u$ *of the solution* u *to* (9) *replacing the* H^1*-norm by the* $W^{1,1}$*-norm. This is due to the similar decay behaviour of the solution's derivatives and the derivatives of its powers. More precisely, for fixed* p*, the decay rate in the estimate* (14) *on the derivatives of* u *and the decay rate in the estimate* (19) *on the derivatives of the powers of* u *coincide. Hence, we obtain the same convergence result for the moment computation under the same asymptotic decay assumption on the sequence* $\{\gamma_k\}$*. Of course, the constant in the convergence result is then affected by* p*.*

Corollary 4.9. *Let* $f \in L^p(D)$ *for* $p \geqslant 2$*. Under the conditions of Theorem 4.3, the quasi-Monte Carlo quadrature using the first* N *Halton points for approximating the* p*-th moment of the solution* u *to* (12) *provides the error estimate*

$$\big\|(\mathbf{I} - \mathbf{Q}_{\Xi_N})\hat{u}^p\big\|_{W^{1,1}(D)} \lesssim \|f\|_{L^p(D)}^p m N^{-1+\delta}$$

with a constant depending on δ *and* p *but not on* m*.*

5. Numerical results

In this section, we present numerical examples to validate the theoretical findings. To that end, we consider the one-dimensional diffusion problem

$$-\partial_x\big(a(x,\mathbf{y})\partial_x u(x,\mathbf{y})\big) = 1 \text{ in } D = (0,1) \tag{47}$$

with homogenous boundary conditions, i.e. $u(0,\mathbf{y}) = u(1,\mathbf{y}) = 0$. The logarithm of the diffusion coefficient a is given by the Karhunen-Loève expansion

$$\log\big(a(x,\mathbf{y})\big) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\varphi_k(x_k)y_k.$$

Here, the eigenpairs (λ_k, φ_k) are obtained by solving the eigenproblem for the diffusion coefficient's correlation, i.e.

$$(\mathcal{C}\varphi_k)(x) = \int_0^1 \kappa(x,x')\varphi_k(x')\,\mathrm{d}x' = \lambda_k\varphi_k(x), \tag{48}$$

where we assume that this correlation is given by a positive definite function

$$\kappa(x,x') := \int_\Omega \log\big(a(x,\omega)\big)\log\big(a(x',\omega)\big)\,\mathrm{d}\mathbb{P}(\omega).$$

The knowledge of $\kappa(x, x')$ together with $\mathbb{E}[a(x,\omega)] = 0$ provides the unique description of a since the underlying random process is Gaussian.

Let $r = |x - x'|$. In the sequel, we consider the class of Matérn correlation kernels, i.e.

$$\kappa_\nu(r) := \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu} r}{\ell} \right)^\nu K_\nu \left(\frac{\sqrt{2\nu} r}{\ell} \right)$$

with $\ell, \nu \in (0, \infty)$. Here, K_ν denotes the *modified Bessel function of the second kind*, cf. [1]. For half integer values of ν, i.e. $\nu = p + 1/2$ for $p \in \mathbb{N}$, this expression simplifies to

$$\kappa_{p+1/2}(r) = \exp\left(\frac{-\sqrt{2\nu} r}{\ell} \right) \frac{p!}{(2p)!} \sum_{i=0}^{p} \frac{(p+i)!}{i!(p-i)!} \left(\frac{\sqrt{8\nu} r}{\ell} \right)^{p-i}.$$

In the limit case $\nu \to \infty$, we obtain the Gaussian correlation

$$\kappa_\infty(r) = \exp\left(\frac{-r^2}{2\ell^2} \right),$$

cf. [39]. The Sobolev smoothness of the kernel κ_ν is controlled by the *smoothness parameter* ν. A visualization of these kernels for different values of ν is given in Figure 1. The eigenvalues of the

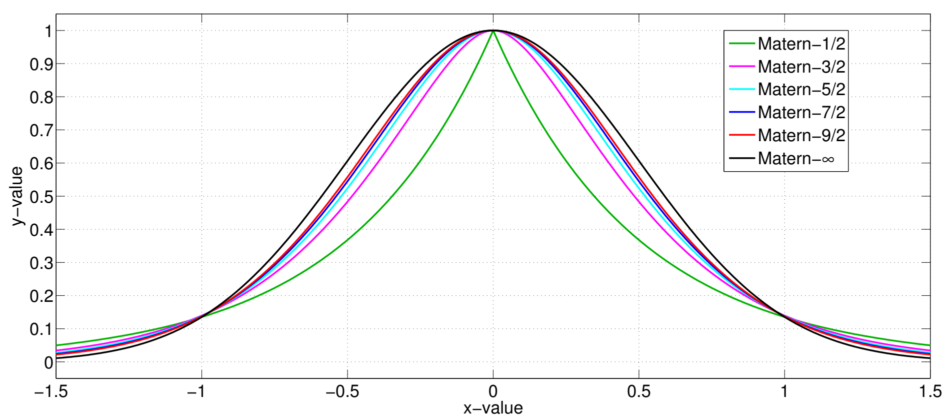

FIGURE 1. Different values for the smoothness parameter ν.

Matérn correlation kernels decay like

$$\lambda_k \leqslant C k^{-(1+2\nu/d)}$$

for some $C > 0$, cf. [15]. Moreover, it has been shown in [15] that the corresponding eigenfunctions φ_k fulfill $\|\varphi_k\|_{L^\infty(D)} \leqslant C k^{1/2+\eta}$ for some $C > 0$ and arbitrary $\eta > 0$. Even so, at least in one dimension, the numerical experiments in [15] suggest that φ_k is bounded independently of k. In this case, for $d = 1$, this leads to the decay

$$\gamma_k \leqslant C k^{-(1/2+\nu)}$$

for the sequence $\{\gamma_k\}_k$.

In the sequel, we consider $\nu = 5/2, 7/2, 9/2$. For the parameter value $\nu = 5/2$, the eigenvalues of the correlation function decay too slowly and we are thus outside our regime. The parameter value $\nu = 7/2$ is exactly the limit case for the decay of the eigenvalues and the value $\nu = 9/2$ leads to an eigenvalue decay perfectly fitting our assumptions. The correlation length is set to $\ell = 1/2$ for each of the kernels. The decay of the related sequences $\{\gamma_k\}_k$ is depicted in Figure 2. As can be seen, we observe for all choices of ν a certain offset before the asymptotical rate is achieved, which is caused by the correlation length.

We have discretized (47) by piecewise linear finite elements and chose piecewise constant elements for the discretization of the diffusion coefficient. A reference solution is computed by the

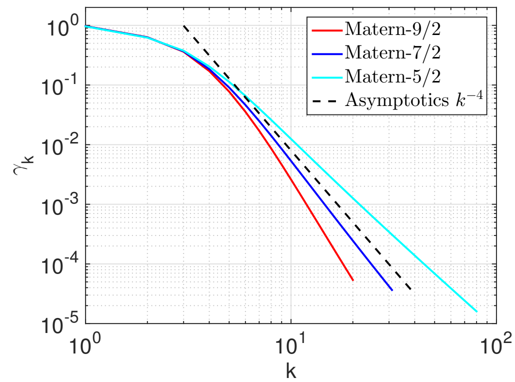

FIGURE 2. Decay of the γ_k for the Matérn kernels under consideration.

quasi-Monte Carlo quadrature with Halton points and $N = 10 \cdot 2^{20} \approx 10^7$ samples. We compute each of these samples by solving (47) on the spatial discretization level 14, i.e. we have the meshwidth $h = 2^{-14}$.

The computation of the truncated Karhunen-Loève expansion is performed by the pivoted Cholesky decomposition of the associated covariance operator $\mathcal{C}$, cf. (48). We choose m in such a way that the trace error in the covariance operator $\mathcal{C}$ is smaller than h^2 in order to rule out the truncation error, see [20] for the details.

The computations for the approximation error are also performed on level 14. This means, we have kept the level fixed and successively increased the number of Halton points.

Since the reference solution is obtained by oversampling, we have to rule out systematic errors. Thus, we think it is appropriate to validate the reference solution by a different method. To that end and also for comparison, we provide the convergence of a Monte Carlo quadrature towards the quasi-Monte Carlo reference. Our numerical realization of the Monte Carlo quadrature is obtained by choosing the sample points $\boldsymbol{\xi}^i$ in (25) as pseudo random numbers which are generated by the *Mersenne Twister*, cf. [30]. The related error plots show the approximation of the root mean square error by averaging five runs of the Monte Carlo quadrature. Beside this additional cost, the computational cost of the Monte Carlo quadrature and the quasi-Monte Carlo quadrature behave rather similar, i.e. the time for computing a single sample by solving the related diffusion problem is the same for both methods. Note that the cost for the generation of the sample points are negligible for both methods.

The Matérn kernel for $\nu = 9/2$. For the smoothness parameter $\nu = 9/2$, which perfectly fits our smoothness assumptions, we have truncated the Karhunen-Loève expansion after $m = 20$ terms. The plots in Figure 3 show the Monte Carlo quadrature errors (left plot) and the quasi-Monte Carlo quadrature errors (right plot) for the mean and the moments up to order 4 with respect to increasing numbers N of sampling points. The Monte Carlo method convergences towards the reference solution at the expected rate $N^{-1/2}$. For the quasi-Monte Carlo quadrature, we obtain convergence rates that are significantly higher than $N^{-1/2}$, at least for the mean and the second moment. The particular rates are indicated by the slopes which can be found in the plot and which corresponds to a linear least-squares fit for the respective curve. The successive decrease of the convergence rate for the higher moments can be explained by the exponential dependence of the constants in (19) on p in the pre-asymptotic regime.

The Matérn kernel for $\nu = 7/2$. For the smoothness parameter $\nu = 7/2$, the Karhunen-Loève expansion has been truncated after $m = 31$ terms. As already mentioned, the underlying Matérn kernel $\kappa_{7/2}$ complies with the limit case for the required smoothness of the correlation kernel. The quadrature errors are found in Figure 4 for the Monte Carlo method (left plot) and for the quasi-Monte Carlo method (right plot), again for the mean and the moments up to order 4 with

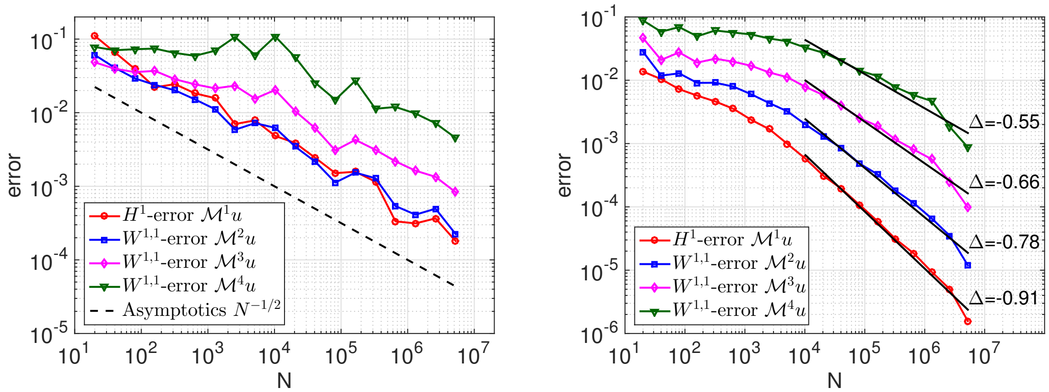

FIGURE 3. Convergence of the Monte Carlo quadrature (left) and the quasi-Monte Carlo reference (right) for $\nu = 9/2$.

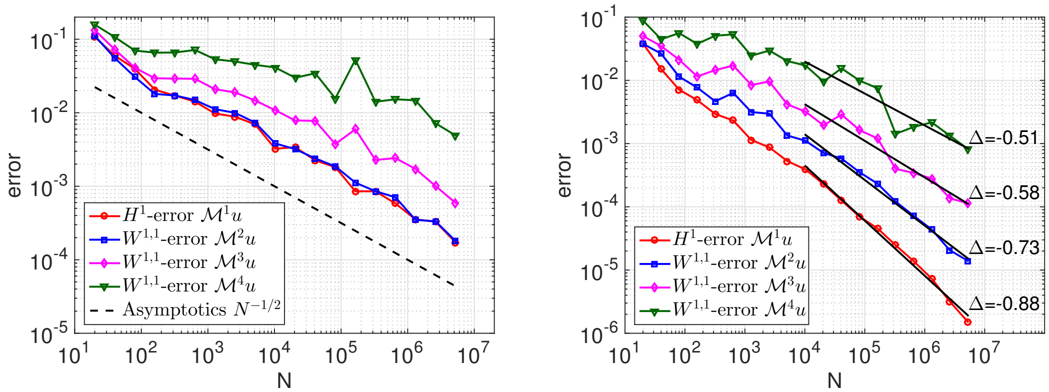

FIGURE 4. Convergence of the Monte Carlo quadrature (left) and the quasi-Monte Carlo reference (right) for $\nu = 7/2$.

respect to increasing numbers N of sampling points. We observe almost the same decay of the errors as in the previous example with the smoothness parameter $\nu = 9/2$.

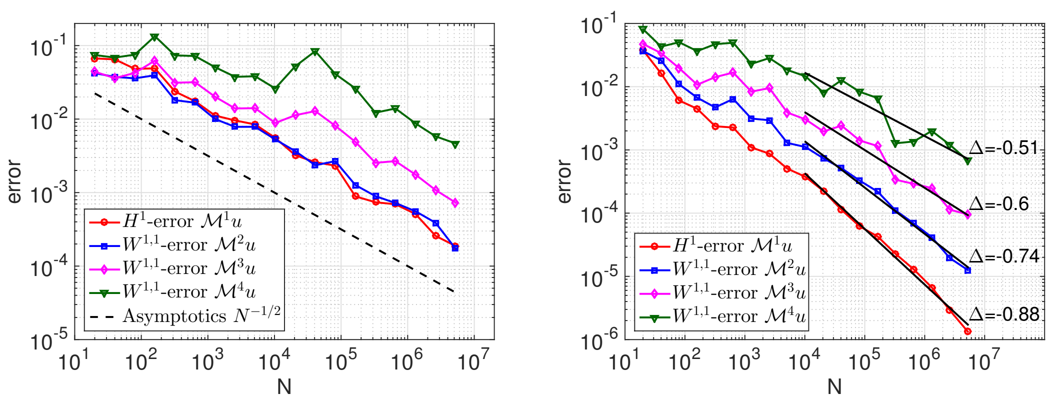

FIGURE 5. Convergence of the Monte Carlo quadrature (left) and the quasi-Monte Carlo reference (right) for $\nu = 5/2$.

The Matérn kernel for $\nu = 5/2$. For the smoothness parameter $\nu = 5/2$, we have truncated the Karhunen-Loève expansion after $m = 80$ terms. A visualization of the corresponding errors for an increasing number of Halton points is given on the right of Figure 5. The convergence of the Monte Carlo quadrature is provided on the left of this figure. Although, the correlation kernel $\kappa_{5/2}$ does not match the required smoothness assumptions anymore, we essentially observe the same error rates as in the previous two examples. Note that a similar effect has been observed in [15] for randomly shifted lattice rules. As mentioned in the introduction, these quadrature rules are well suited for the problem at hand and the error estimates require less regularity, i.e. they converge with a rate $N^{-1+\delta}$ in the mean square sense whenever $\gamma_k \lesssim k^{-3/2-\eta}$. The numerical examples in [15] consider a decay $\gamma_k \approx k^{-2}$ and $\gamma_k \approx k^{-5/4}$ and both examples yield nearly the same convergence rates.

We may summarize the numerical results as follows. In our examples, we observe essentially the same convergence behavior of the quasi-Monte Carlo quadrature based on Halton points with respect to the smoothness parameter ν of the Matérn covariance kernels. This smoothness parameter determines both, the stochastic dimension m, given by the length of the Karhunen-Loève expansion, and also the decay of the sequence $\{\gamma_k\}$. This indicates, in concordance with Theorem 4.3 and Remark 4.5, convergence rates for the quasi-Monte Carlo quadrature based on Halton points which are almost independent of the smoothness of the underlying covariance kernel. The results also imply that the claimed smoothness assumptions, i.e. $\gamma_k \lesssim k^{-4-2\eta}$, can probably be weakened.

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, Applied mathematics series, Dover Publications, New York, 1964.

[2] R. A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[3] E. Atanassov, *On the discrepancy of the Halton sequences.*, Math. Balkanica (N.S.), 18 (2004), pp. 15–32.

[4] I. Babuška, F. Nobile, and R. Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM J. Numer. Anal., 45 (2007), pp. 1005–1034.

[5] I. Babuška, R. Tempone, and G. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM J. Numer. Anal., 42 (2004), pp. 800–825.

[6] A. Barth, C. Schwab, and N. Zollinger, *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*, Numer. Math., 119 (2011), pp. 123–161.

[7] J. Charrier, *Strong and weak error estimates for elliptic partial differential equations with random coefficients*, SIAM J. Numer. Anal., 50 (2012), pp. 216–246.

[8] J. Charrier, R. Scheichl, and A. Teckentrup, *Finite element error analysis of elliptic pdes with random coefficients and its application to multilevel Monte Carlo methods*, SIAM J. Numer. Anal., 51 (2013), pp. 322–352.

[9] G. M. Constantine and T. H. Savits, *A multivariate Faà di Bruno formula with applications*, Trans. Amer. Math. Soc., 248 (1996), pp. 503–520.

[10] P. Fan, *New inequalities of Mill's ratio and its application to the inverse Q-function approximation*, Aust. J. Math. Anal. Appl., 10 (2013), p. Art. 11.

[11] P. Frauenfelder, C. Schwab, and R. Todor, *Finite elements for elliptic problems with stochastic coefficients*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 205–228.

[12] R. Ghanem and P. Spanos, *Stochastic Finite Elements: a Spectral Approach*, Springer, New York, 1991.

[13] M. Giles, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617.

[14] C. J. Gittelson, *Stochastic Galerkin discretization of the log-normal isotropic diffusion problem.*, Math. Models Methods Appl. Sci., 20 (2010), pp. 237–263.

[15] I. G. Graham, F. Y. Kuo, J. A. Nichols, R. Scheichl, C. Schwab, and I. H. Sloan, *Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients*, Numer. Math., (2015). to appear.

[16] J. H. Halton, *On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals*, Numer. Math., 2 (1960), pp. 84–90.

[17] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*, Methuen, London, 1964.

[18] H. Harbrecht, M. Peters, and M. Siebenmorgen, *Multilevel accelerated quadrature for PDEs with log-normal distributed random coefficient*, Preprint 2013-18, Institute of Mathematics, University of Basel, (2013).

[19] ———, *On multilevel quadrature for elliptic stochastic partial differential equations*, in Sparse Grids and Applications, J. Garcke and M. Griebel, eds., vol. 88 of Lecture Notes in Computational Science and Engineering, Springer, Berlin-Heidelberg, 2013, pp. 161–179.

[20] ———, *Efficient approximation of random fields for numerical applications*, Numer. Linear Algebra Appl., (2015). to appear.

[21] S. Heinrich, *The multilevel method of dependent tests*, in Advances in stochastic simulation methods (St. Petersburg, 1998), Stat. Ind. Technol., Birkhäuser, Boston, MA, 2000, pp. 47–61.
[22] ———, *Multilevel Monte Carlo methods*, in Lecture Notes in Large Scale Scientific Computing, London, 2001, Springer, pp. 58–67.
[23] F. J. Hickernell, *A generalized discrepancy and quadrature error bound*, Math. Comp., 67 (1998), pp. 299–322.
[24] F. J. Hickernell, I. H. Sloan, and G. W. Wasilkowski, *On tractability of weighted integration over bounded and unbounded regions in* $\mathbb{R}^s$, Math. Comp., 73 (2004), pp. 1885–1901 (electronic).
[25] E. Hlawka, *Über die Diskrepanz mehrdimensionaler Folgen mod. 1*, Mathematische Zeitschrift, 77 (1961), pp. 273–284.
[26] V. A. Hoang and C. Schwab, *N-term Wiener chaos approximation rates for elliptic PDEs with lognormal gaussian random inputs*, Math. Models Methods Appl., 24 (2014), pp. 797–826.
[27] F. Kuo, C. Schwab, and I. Sloan, *Quasi-Monte Carlo methods for high-dimensional integration: the standard (weighted Hilbert space) setting and beyond*, ANZIAM Journal, 53 (2012), pp. 1–37.
[28] F. Kuo, G. Wasilkowski, and B. Waterhouse, *Randomly shifted lattice rules for unbounded integrands*, J. Complexity, 22 (2006), pp. 630 – 651.
[29] M. Loève, *Probability Theory. I+II*, vol. 45 of Graduate Texts in Mathematics, Springer, Berlin-Heidelberg-New York, 4th ed., 1977.
[30] M. Matsumoto and T. Nishimura, *Mersenne Twister: a 623-dimensionally equidistributed uniform pseudo-random number generator*, ACM Trans. Model. Comput. Simul., 8 (1998), pp. 3–30.
[31] H. Matthies and A. Keese, *Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 1295–1331.
[32] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.
[33] F. Nobile, R. Tempone, and C. G. Webster, *An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal., 46 (2008), pp. 2411–2442.
[34] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems: Linear Information*, EMS Tracts in Mathematics, European Mathematical Society, Zürich, 2008.
[35] ———, *Tractability of Multivariate Problems: Standard Information for Functionals*, EMS Tracts in Mathematics, European Mathematical Society, Zürich, 2010.
[36] ———, *Tractability of Multivariate Problems: Standard Information for Operators*, EMS Tracts in Mathematics, European Mathematical Society, Zürich, 2012.
[37] A. B. Owen, *Halton sequences avoid the origin*, SIAM Rev., 48 (2006), pp. 487–503.
[38] J. K. Patel and C. B. Read, *Handbook of the Normal Distribution*, CRC Press, Boca Raton, 1996.
[39] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*, MIT Press, Cambridge, 2005.
[40] J. Sándor, D. S. Mitrinovic, and B. Crstici, *Handbook of Number Theory I*, Mathematics and Its Applications, Springer, Berlin Heidelberg New York, 1995.
[41] C. Schwab and C. J. Gittelson, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, Acta Numer., 20 (2011), pp. 291–467.
[42] C. G. Simader, *On Dirichlet's Boundary Value Problem: LP-Theory based on a Generalization of Garding's Inequality*, Lecture Notes in Mathematics, Springer, Berlin-Heidelberg-New York, 1972.
[43] I. H. Sloan, F. Y. Kuo, and S. Joe, *Constructing randomly shifted lattice rules in weighted Sobolev spaces*, SIAM J. Numer. Anal., 40 (2002), pp. 1650–1665.
[44] I. H. Sloan and H. Woźniakowski, *When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?*, J. Complexity, 14 (1997), pp. 1–33.
[45] X. Wang, *A constructive approach to strong tractability using quasi-Monte Carlo algorithms*, J. Complexity, 18 (2002), pp. 683–701.
[46] H. Woźniakowski, *Tractability and strong tractability of linear multivariate problems*, J. Complexity, 10 (1994), pp. 96–128.
[47] S. K. Zaremba, *Some applications of multidimensional integration by parts*, Ann. Polon. Math., 21 (1968), pp. 85–96.

Helmut Harbrecht, Michael Peters, Markus Siebenmorgen, Departement Mathematik und Informatik, Universität Basel, Spiegelgasse 1, 4051 Basel, Schweiz

E-mail address: {helmut.harbrecht,michael.peters,markus.siebenmorgen}@unibas.ch